

HAUTE AUTORITÉ DE SANTÉ

Assessing the Readability of Clinical Guidelines with Deontic Markers


¹Gersende Georg, ²Mark Truran,
³Dong Zhou, ²Marc Cavazza

¹French National Authority for Health, Paris, France;
²School of Computing, Teesside University, United Kingdom;
³Dept. of Computer Science, Trinity College, Dublin, Ireland

Guidelines International Network 2011 - 30/08/2011

Introduction

- **Research in guidelines *quality* has mostly been based on ‘user’ evaluation rather than formal properties of guidelines as textual documents**
- **We introduce the use of *readability metrics* and adapt them to the specific context of clinical guidelines**



Guidelines International Network 2011 - 30/08/2011

Key Ideas

- Explore the readability of clinical guidelines
- Compare the use of generic readability metrics to guidelines-specific ones (based on previous work on guidelines' linguistic properties)



Guidelines International Network 2011 - 30/08/2011

Standard Readability measures

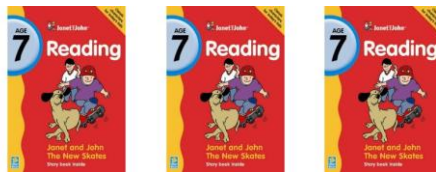
- Readability indices *approximate* the reading difficulty of a text
- They have been used extensively by educators and librarians since the 1920s¹
- These indices exploit the *surface or explicit* features of text
 - Sentence lengths
 - Word frequencies



¹Vogel et al. An objective method of determining grade placement of children's reading material. *Elementary School Journal*, 28:373-381, 1928.

From generic readability measures to specialised ones

- **However, there is a problem - readability indexes were designed with *grade school* literature in mind**
 - At this task they are surprisingly effective
- **But previous work has shown that lexical formulae alone are not always effective when applied to highly technical texts**



HAS

Guidelines International Network 2011 - 30/08/2011

Linguistic properties of clinical guidelines (Georg et al. MIE 2005)

- **Specific rhetorical structures associated to the expression of recommendations: Deontic Operators**
- **Some sections are similar in content to generic texts; others are more specialised in their use of terminology**
- **Deontic Operators play a role in structuring recommendations**

¹Moulin B, Rousseau D. Knowledge acquisition from prescriptive texts. ACM, 1990: 1112:1121.

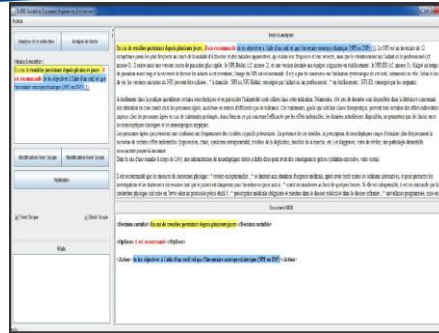
²Georg G, Colombet I, Jaulent MC. Structuring Clinical Guidelines through the Recognition of Deontic Operators. Stud Health Technol Inform. 2005;116:151-6.

³Georg G, Pauchet-Traversat AF, André-Vert J, Geffrier-d'Acremont C. Integrating a document engineering environment into the French guideline development process. G-I-N Conference 2007. Toronto, Ontario, Canada. 22-25 August 2007.

HAS

G-DEE for Guidelines Document Engineering Environment

- We have developed an intelligent document engineering environment (G-DEE) to assist with guideline production
- It assesses the 'quality' of guidelines by automatically analysing their structure based on the recognition of key recommendations
- These recommendations are detected using shallow NLP technique which recognise specific linguistic structures known as deontic operators¹
- G-DEE has been successful deployed at HAS since 2007 being spontaneously adopted by a majority of project leaders



¹Georg G. Jaulent MC. An Environment for Document Engineering of Clinical Guidelines. In: CP Friedman, editor. Proceedings of the American Medical Informatics Association; 22-26 October 2005; Washington, DC, 2005:276-280.

VI. TRAITEMENT PHARMACOLOGIQUE (150).

VI.1. Choix des antihypertenseurs (151).

La réduction du risque cardio-vasculaire est avant tout dépendante de la baisse de la pression artérielle, quelle que soit la classe d'antihypertenseur utilisée (grade A) (86).

Dans l'HTA essentielle non compliquée, les 5 classes d'antihypertenseurs majeurs [les diurétiques thiazidiques, les bêta-bloquants, les inhibiteurs calciques, les inhibiteurs de l'enzyme de conversion (IEC) et les antagonistes des récepteurs de l'angiotensine II (ARAII)] ont montré un bénéfice sur la morbi-mortalité cardio-vasculaire, dans les essais cliniques (87).

Ces 5 classes d'antihypertenseurs peuvent donc être proposées en première intention dans la prise en charge d'un hypertendu essentiel non compliqué (grade A) (88). Le choix d'un traitement médicamenteux sera adapté à chaque patient en fonction des indications préférentielles de certaines classes dans des situations cliniques articulées (cf (89).

infra) (en accord avec les études cliniques), de l'efficacité et de la tolérance des médicaments déjà pris par le patient, de l'existence de comorbidités pouvant justifier ou contre-indiquer certains antihypertenseurs, et du coût du traitement et de sa surveillance, en sachant que le diurétique thiazidique fait partie des classes dont le coût journalier est le plus faible (90).

En cas de prescription d'un traitement par IEC ou ARA-II, il est recommandé de prescrire un bilan associant kaliémie et créatininémie dans un délai de 7 à 15 jours, après le début du traitement (91).

Si la créatininémie s'élève de plus de 20 à 30 %, il est recommandé d'arrêter l'IEC ou l'ARA-II et de demander un avis spécialisé (92). Pour favoriser l'observance, une prise unique quotidienne (monoprise) utilisant un médicament de longue durée d'action sera préférée (93).

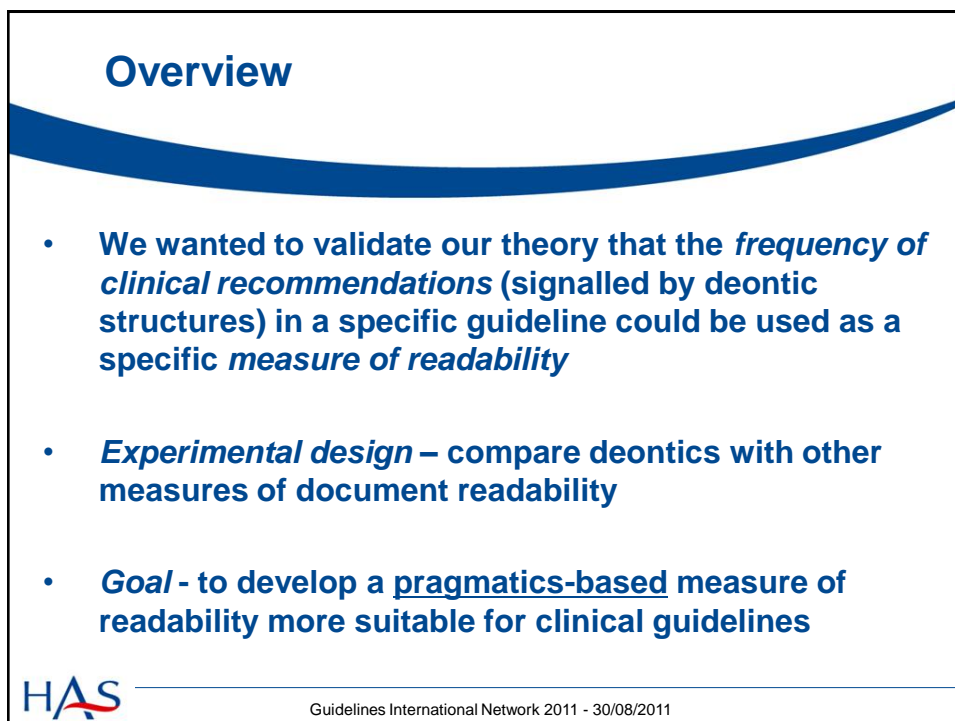
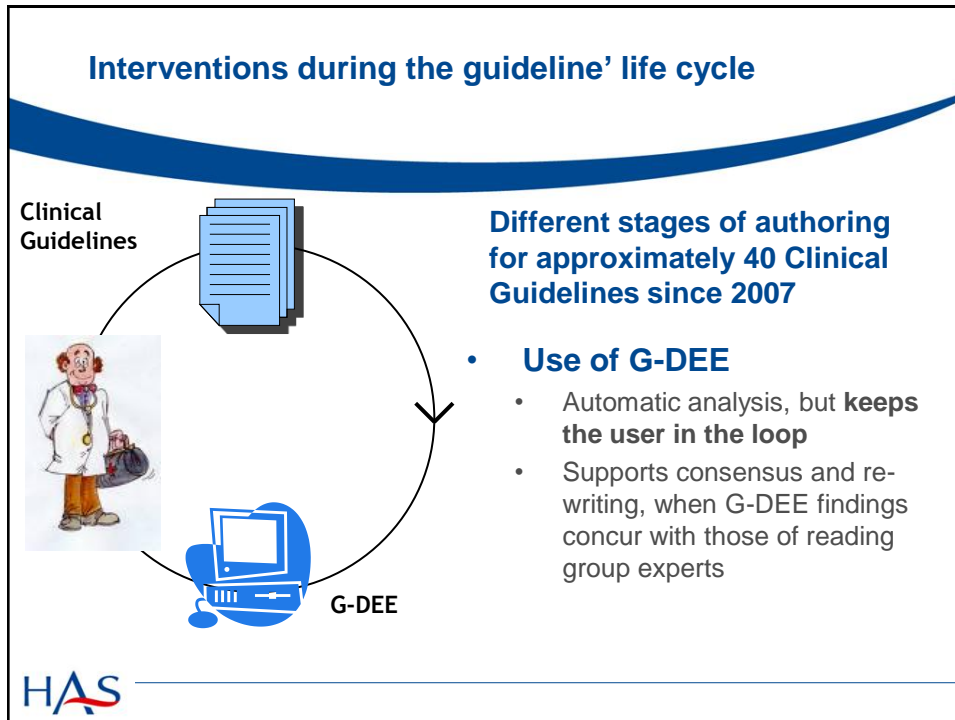
A partir d'études menées chez des patients afro-américains et africains des Caraïbes, il apparaît que l'HTA chez ces patients, est plus sensible aux diurétiques et aux inhibiteurs calciques qu'aux classes thérapeutiques agissant sur le système rénine-angiotensine (IEC, ARA-II) ou qu'aux bêta-bloquants (94).

VI.2. Stratégie d'adaptation du traitement médicamenteux (152).

Il est recommandé de débiter par une monothérapie (95). Une association fixe d'antihypertenseurs à doses faibles, ayant l'AMM en première intention pour l'indication HTA, peut également être proposée (96).

En deuxième intention, une bithérapie sera instaurée dans un délai d'au moins 4 semaines, en cas de réponse tensionnelle insuffisante au traitement initial (97).

Une bithérapie pourra être instaurée dans un délai plus court, dans les cas suivants chez le patient ayant une PA = 180-110 mmHg quel que soit le nombre de facteurs de risque cardio-vasculaire associés, chez le patient ayant une PA de 140-179/90-109 mmHg et à risque cardio-vasculaire élevé (98).



New Readability Metric: Normalised Deontic Frequency

- We processed 10 (French) guidelines published by HAS using G-DEE
- We calculated the normalised deontic frequency *NDF* as

$$NDF = \frac{DSW}{L}$$

- Where *DSW* is the total number of words within deontic structures and *L* is the total number of words in the document



Guidelines International Network 2011 - 30/08/2011



Calculation of the Flesch reading Ease Index (non-specific)

- We then scored each clinical guideline using the Flesch Reading Ease index for French texts

$$FRE = \frac{209 - (0.68 \times SW) - (1.15 \times SL)}{100}$$

- Where *SW* is the average number of syllables per word and *SL* is the average sentence length (in words)
- This produced a value in the range [0,1] with a score of 1 indicating a document that is very easy to read



Guidelines International Network 2011 - 30/08/2011



Latent Semantic Analysis (non-specific, 2)

- **Latent Semantic Analysis (LSA) is a statistical technique which gives access to semantic properties of text without the need to define explicit semantic categories – these are automatically derived from training corpora using statistical IR techniques (single value decomposition of lexical matrices)**
- **It has been shown to provide a reliable measure of ‘textual cohesion’, which also correlates to readability**
- **It is applied on individual text sections which are centred on a limited number of topics**



Guidelines International Network 2011 - 30/08/2011

Calculation of the Latent Semantic Analysis

- **We built the semantic space using the ALPESMED corpus (University of Grenoble)**
- **We selected 60 PDF tutorials at random from the corpus and converted them to plain text (63398 terms)**
- **We built the semantic space using a combination of the Terrier IR platform and the JAMA package**
 - The processing took 6 hours, producing a 900MB semantic space (a text file) – needed only once



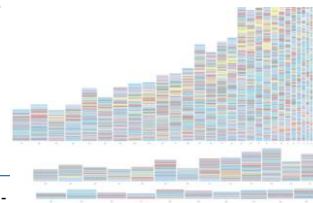
Guidelines International Network 2011 - 30/08/2011

Measurement of the Textual Cohesion (COH)

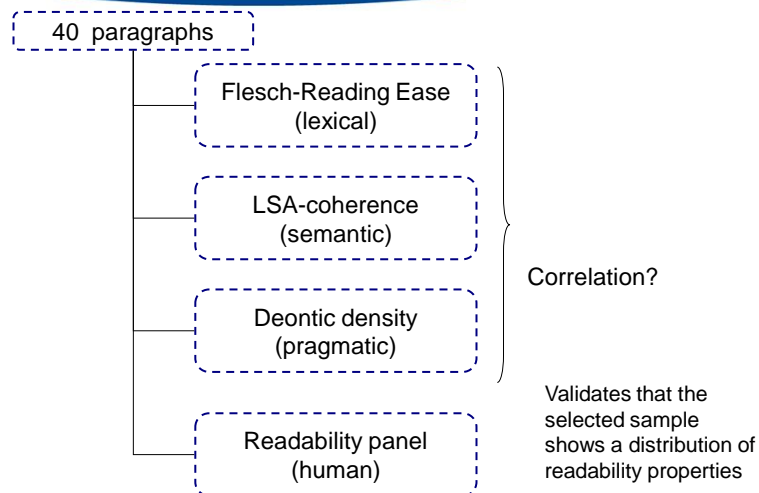
- We then measured the textual cohesion of each clinical guideline in turn
- We did this by calculating the textual cohesion of all pairs of adjacent sentences so that

$$COH = \frac{\sum_{i=1}^{n-1} COS_{i,i+1}}{n-1}$$

- This yielded a score for each range [-1,1]



Overview of the experiment



The experiment

- We extracted a sample of 40 paragraphs from the guidelines
- We scored each paragraph as above, and we also passed each paragraph to a *panel of domain experts*
- The panel assigned each paragraph a readability score using a 5-point Likert scale



Guidelines International Network 2011 - 30/08/2011

Readability Questionnaire

1. The topic under discussion can be identified immediately and the section contents are clear
2. The section is readable but rather tedious
3. The section requires reading effort to be properly understood
4. I had to read some sentences again to understand their meaning
5. Even after reading the section several times, its meaning is not entirely clear



Guidelines International Network 2011 - 30/08/2011

Results (1)

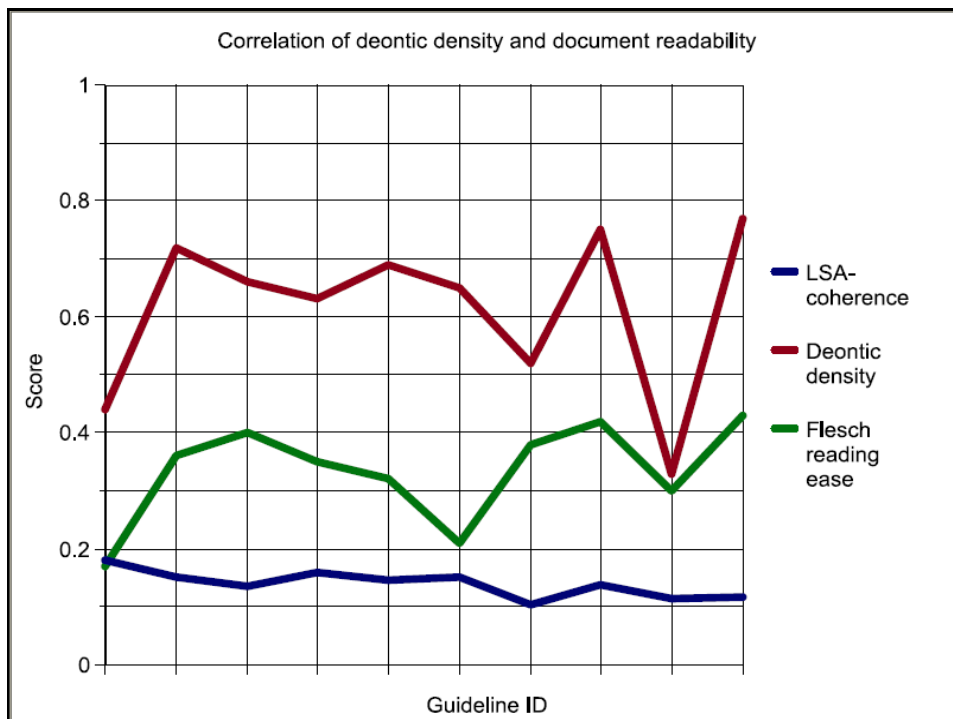
- There was a significant positive correlation between deontic density and Flesch Reading Ease index (suggests a positive impact of deontic structures on the underlying text)
- At first, there was no correlation between LSA-cohesion and deontic density (is the structure or content of guidelines impairing the application of LSA techniques?)

	LSA	NDF	FRE
LSA	-	-0.11 (p-value=0.72)	-0.38 (p-value=0.15)
NDF	-0.11 (p-value=0.72)	-	0.55 (p-value=0.03)
FRE	-0.38 (p-value=0.15)	0.55 (p-value=0.03)	-

Table 2: Kendall tau (τ) correlation coefficients (document-level readability)



Guidelines International Network 2011 - 30/08/2011

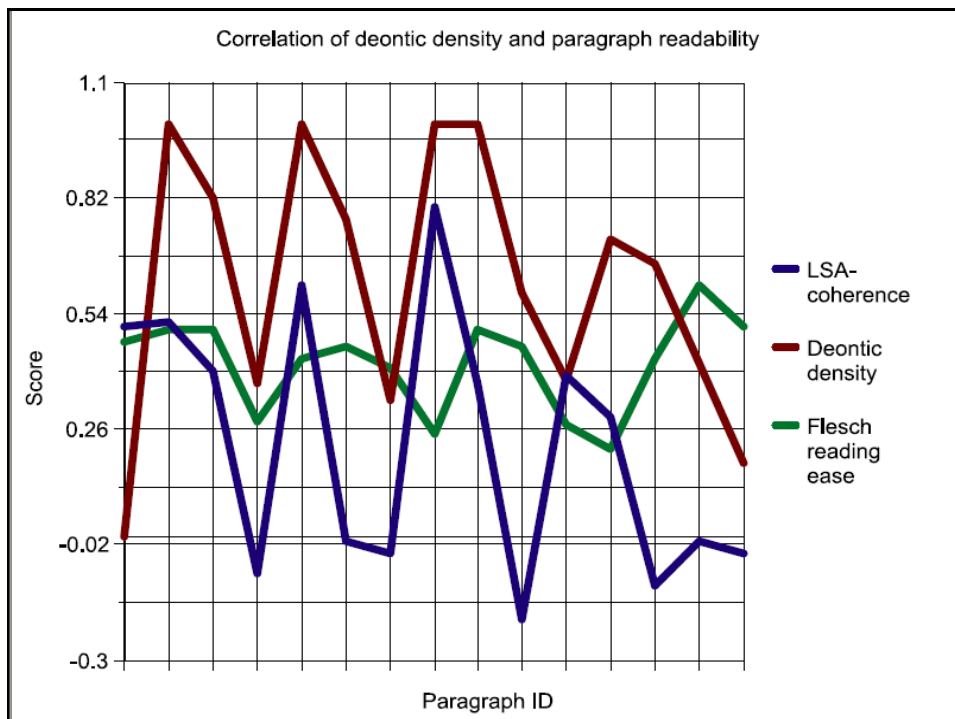


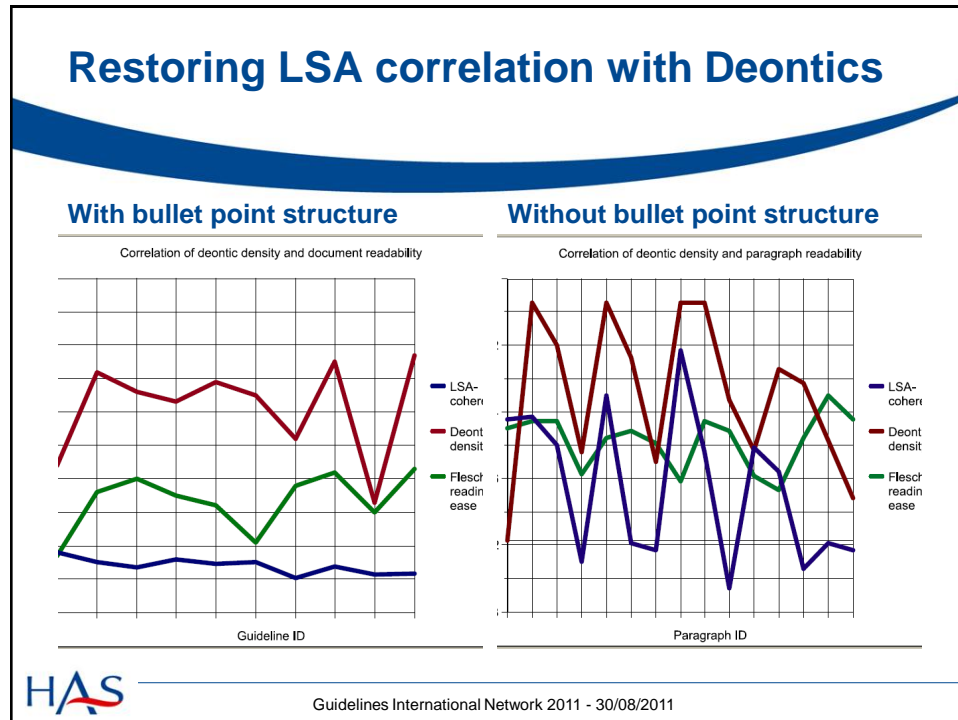
Results (2)

- We realised that high (excessive?) use of bullet point structures impaired the application of LSA analysis (see slide 15, measure of COH is based on adjacent sentences)
- We then adjusted our paragraph selection criterion to exclude excessive bullet point structures, leaving us with 15 'conventional' paragraphs
- The correlation between normalised deontic density and LSA-cohesion jumped to a highly significant $\tau=0.50$ ($p\text{-value}=0.013$)!



Guidelines International Network 2011 - 30/08/2011





Conclusions

- **Until now the only validation of the importance of deontic structures (distribution, scope ratio ...) was empirical, through the HAS experience**
- **These results provide an insight into the actual impact of deontic structures on the overall readability of recommendations**
- **HYP: the use of deontic structures has a positive impact on the overall readability of the embedded text (scope) without imposing specific constraints on writing**

HAS Guidelines International Network 2011 - 30/08/2011

For More (Technical) Details:

Truran M, Georg G, Zhou D, Cavazza M. **Assessing the Readability of Clinical Guidelines in a Document Engineering Environment**. In: Apostolos Antonacopoulos and Michael J. Gormish and Rolf Ingold (Eds.) *DocEng'10 - Proceedings of the 10th ACM Symposium on Document Engineering, September 21-24, 2010, Manchester, UK*. ACM Press, New York NY, USA. 2010;:125-134.

Can be downloaded at: <http://gersende.georg.free.fr>

Contact Information

Dr Gersende GEORG
Haute Autorité de Santé,
Clinical Guidelines Unit & Healthcare Organizations
Accreditation Unit, Paris, France

E-mail: g.georg@has-sante.fr



Guidelines International Network 2011 - 30/08/2011