Expanding the Diversity of Texts and Applications: Findings from the Section on Clinical Natural Language Processing of the International Medical Informatics Association Yearbook

Aurélie Névéol, Pierre Zweigenbaum, Section Editors for the IMIA Yearbook Section on Clinical Natural Language Processing
LIMSI, CNRS, Université Paris-Saclay, Orsay, France

Summary

Objectives: To summarize recent research and present a selection of the best papers published in 2017 in the field of clinical Natural Language Processing (NLP).

Methods: A survey of the literature was performed by the two editors of the NLP section of the International Medical Informatics Association (IMIA) Yearbook. Bibliographic databases PubMed and Association of Computational Linguistics (ACL) Anthology were searched for papers with a focus on NLP efforts applied to clinical texts or aimed at a clinical outcome. A total of 709 papers were automatically ranked and then manually reviewed based on title and abstract. A shortlist of 15 candidate best papers was selected by the section editors and peer-reviewed by independent external reviewers to come to the three best clinical NLP papers for 2017.

Results: Clinical NLP best papers provide a contribution that ranges from methodological studies to the application of research results to practical clinical settings. They draw from text genres as diverse as clinical narratives across hospitals and languages or social media.

Conclusions: Clinical NLP continued to thrive in 2017, with an increasing number of contributions towards applications compared to fundamental methods. Methodological work explores deep learning and system adaptation across language variants. Research results continue to translate into freely available tools and corpora, mainly for the English language.

Keywords

Awards and prizes; medical informatics/trends; natural language processing; semantics

Yearb Med Inform 2018:193-8
http://dx.doi.org/10.1055/s-0038-1667080

Introduction

Clinical Natural Language Processing (NLP) is defined as Natural Language Processing applied to clinical texts or aimed at a clinical outcome. This encompasses NLP applied to texts in Electronic Health Records (EHRs), which is the case of the bulk of information extraction for decision support or clinical research. We also consider as clinically relevant, applications and research addressing the analysis of patient-authored text or speech for public health or diagnosis purposes. As noted in 2017, some NLP tools and methods are freely available and easy to use for researchers in other fields to apply routinely to their application needs. As a result, some studies using NLP may appear as best papers in other sections of the IMIA Yearbook. However, their selection criteria focus on the different section topics rather than emphasize the NLP methodology.

This year's survey paper reports on the contribution of shared tasks to the advancement of clinical NLP research [1]. Best papers selected this year provide contributions that each cover a variety of texts of clinical interest: Pérez et al. [2] address medical named entity recognition in clinical text in Spanish and Swedish; furthermore, they emphasize methods' contribution in a context where little training data is available, which is often the case for languages other than English or when a new medical specialty is explored. Tapi Nzali et al. [3] analyze the content of several breast cancer social media venues in French. Castro et al. [4] tackle the extraction of mammography information from radiology reports in English from 18 hospitals. Best papers also cover the continuum between foundational methods and applications. The first paper [2] provides an in-depth methodological focus on named entity recognition. The second paper [3] characterizes breast cancer patients' use of social media and links salient discussion topics to clinical questionnaires with the perspective of recommending updates of the questionnaires, which are important clinical tools for cancer patient management. Finally, the third paper [4] studies information extraction from an application point of view and evaluates symbolic and statistical methods to assess how well the extraction task may generalize across hospitals.

About the Selection Process

Papers were retrieved relying on PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and the Association of Computational Linguistics (ACL) Anthology Searchbench (http://aclasb.dfki.de/). This year, two PubMed queries using free-text keywords and minimal metadata were used sequentially (Figure 1). We do not rely on MeSH terms in the queries because at the time of retrieving articles for review, many have not been indexed yet.

The ACL Anthology query restricted our selection to the most selective journals (Transactions of the Association for Computational Linguistics (TACL), Computational Linguistics), conferences (ACL, Empirical Methods in NLP (EMNLP), North

```
[1] (English[LA] AND journal article[PT] AND 2017[dp] AND
hasabstract[text]) AND ((medical OR clinical OR natural) AND
"language processing").
[2] (English[LA] AND journal article[PT] AND 2017[dp] AND
hasabstract[text]) AND ("text mining") NOT [1].
```

Fig. 1 PubMed queries: [1] PubMed NLP query; [2] PubMed Text Mining query.

American Chapter of the ACL (NAACL), Computational Linguistics (COLING)) and workshops (ACL BioNLP). It used the free text keywords "medical", "clinical", and "health". The collection of papers using these queries brought back 414 titles and abstracts from the PubMed NLP query, 224 additional titles and abstracts from the PubMed Text Mining query and 71 from the ACL Anthology, resulting in a total selection of 709 papers.

Then, we used the results of the 2017's selection to train a logistic regression classifier to automatically rank the selected papers from most relevant to less relevant [5]. The section editors then used the BibReview tool (https://pypi.python.org/pypi/BibReview) to classify the papers based on titles and abstracts into four categories: (1) Off Topic (OT) which identified papers that focused on topics outside of the scope of clinical NLP, such as biological natural language processing, knowledge representation, psycholinguistics, or even image processing; (2) No (N) for papers that did not provide a contribution to either NLP methodology or clinical outcome. Review papers and correspondences were included in this category in order to keep only original research contributions; (3) Maybe (M) for papers that offered a contribution to NLP aimed at a clinical outcome; and (4) Yes (Y) for papers that did so outstandingly, or with high novelty. At the end of this process, the inter-annotator agreement was 85% (600 identical assessments over the 709 papers). Most of the differences involved the category "Maybe". The section editors then reviewed the papers that were rated as "Yes" by at least one editor to arrive at the selection of 15 candidate best papers. We also checked the proportion of OT papers in the set of papers retrieved by each of the three queries used this year. We found that 34% of papers were "Off Topic" in the ACL anthology set, 35% in the Pubmed NLP set and 60% in the Pubmed Text Mining set. This is notable for two reasons. First, it illustrates the difficulty of crafting queries that will efficiently retrieve all and only the clinical NLP literature. Second, it also suggests that clinical NLP work is conducted by several communities who have their own way of describing this type of research.

Results

We present an overview of clinical NLP publications that cover the topics addressed by the research community in 2017. The publications listed with a star (*) in the list of references provided at the end of this synopsis indicate the articles that were reviewed as candidate best papers. A summary of the three best papers appears in the Appendix of this synopsis.

We noticed that 2017 publications focused less on the development of new methods and more on the application of existing methods to unexplored types of texts (including languages other than English, which were addressed by 4.3% of the papers overall) and clinical goals. To group these publications into related subsets, we considered multiple dimensions: the sources they address (social media, patient speech, electronic health records, news, etc.), their authors (consumers, health care professionals, etc.), the NLP tasks performed on these sources (information extraction, classification, etc.), and whether the publication emphasized applications or methods. In this overview of the 2017 literature, we describe the main trends and illustrate them using representative articles, mainly selected from our list of candidate best papers.

Natural Language Processing of Informal Text and Patient-authored Text

This work is well covered by the web mining community and often appears in venues such as journals of the Journal of Medical Internet Research (JMIR) series. As emphasized in Gonzalez et al. [6], social media has become an important alternative source of health-related information on patients. A number of publications address mental health, focusing on disorders such as depression, schizophrenia, dementia, among others. Cheng et al. [7] performed a web-based survey among Chinese microblog users to assess five suicide risk factors. Logistic regression determined the main associations of these risk factors to keywords found in the Chinese version of a standard dictionary for this type of studies, the Linguistic Inquiry and Word Count Lexicon. It is one of the few such studies performed on Chinese so far, and it revealed some strong associations of language markers to mental states such as a high suicide probability marked by a higher usage of pronouns.

Specific analyses, often based upon topic models, are performed to determine the main themes addressed by social media users. Among these, Miller et al. [8] classified tweets about Zika. A first-stage supervised classifier determined Zika-related tweets based on unigram features, then a second-stage classifier divided them into four dimensions of Zika: symptoms, transmission, treatment, and prevention. Then topic analysis with Latent Dirichlet Allocation (LDA) determined the five main topics mentioned for each of the four dimensions. This type of content analysis tool is expected to help disease surveillance research. Tapi Nzali et al. [3] also used LDA to detect the topics in breast cancer posts on Facebook and on a forum. They aligned them with those of the internationally standardized self-administered questionnaires used in cancer clinical trials. They obtained a good alignment but evidenced additional topics

that might help complete current questionnaires. This paper is further described in the Appendix. As another example of this type of research, Lu et al. [9] performed clustering to detect health care stakeholders groups in lung cancer, diabetes, and breast cancer forums. They then applied a second level of clustering on word n-grams and Unified Medical Language System (UMLS) concepts obtained through MetaMap to collect the main topics in these messages. Sentiment analysis was finally performed using the SentiWordNet lexicon. Hao et al. [10] used LDA topic modeling to compare the topics in positive and negative patient reviews in two doctor rating websites in the United States and China. What makes this application of NLP original is its cross-country comparison. Some prominent words in the elicited topics are similar across countries, while others reflect cultural and organizational differences in the two countries and their health care systems, which might have applications for improving health care services. Amith et al. [11] used distributed vector representations learnt from large bodies of text and Pathfinder Network scaling, a technique from cognitive psychology, to model the vaccine-related knowledge structures of health experts and health consumers based upon on-line sources: expert-authored texts intended for a health consumer audience, and consumer-authored USENET group messages. Their comparison revealed knowledge gaps that suggest opportunities to improve provider-patient communication. These examples illustrate how NLP methods can be applied to perform content analysis of online patient-authored text and to collect potentially valuable information and knowledge.

The study of pathological patient speech is well covered by the NLP community and often appears in venues supported by ACL. It continues to be addressed, following up on work that we discussed in the 2016 synopsis, now exploring automated approaches for scoring semantic fluency responses [12].

Natural Language Processing of Text Produced by Health Professionals in the Course of their Medical Practice

This work corresponds to classical topics in clinical NLP, including information extraction from electronic health records (EHRs), patient phenotyping, and EHR classification.

Open source systems are important tools for advancing clinical NLP as they facilitate reproducibility and foster collaborations. Two of the candidate best papers this year specifically focused on making the systems presented available to the community. Kang et al. [13] describe an open source system that analyzes eligibility criteria for clinical trials in the domain of Alzheimer's disease. The system performs the automatic extraction of eligibility criteria from clinical trial descriptions and relies on state-of-the-art methods for entity recognition, negation detection, and relation identification. The study offers an extrinsic evaluation. This is a valuable step in the direction of application-oriented evaluations of NLP systems. However, as pointed out by the authors in their candid limitation section, the system uses simplified queries and the results of the analysis should be confirmed on a more realistic task. Iqbal et al. [14] present the development and evaluation of a multi-phase rule-based pipeline for the recognition and classification of adverse drug effects (ADEs) in psychiatric narratives. The system is open source and includes an improved version of the popular ConText algorithm. The dataset used is multi-centric as it comprises clinical notes from four UK-based psychiatric hospitals. The system shows good performance at detecting ADEs related to antipsychotics and antidepressant drugs. Soysal et al. [15] provide a special attention to system usability with the introduction of CLAMP, a clinical NLP toolkit that implements state-of-the-art NLP methods for entity recognition and linking together with a user-friendly graphic user interface that can help users quickly build customized NLP pipelines for their individual applications.

Method portability is a serious issue, particularly given the variation in "local dialects" found in the free text of clinical narratives produced in different local settings (hospitals) with the same therapeutic goal. Three of the candidate best papers this year are remarkable for addressing the issue of clinical NLP system portability across institutions. Ye et al. [16] investigate the portability of diagnostic systems based on NLP and a Bayesian classifier using clinical notes from two emergency departments to detect influenza cases. They study the causes for performance reduction with an ANOVA analysis and provide a very interesting discussion of what transfer learning is. Sohn et al. [17] use birth cohorts from two US-based hospitals to characterize linguistic variations in reports of asthma in children. The authors use the results of their analysis to explore NLP system portability for asthma ascertainment in the context of retrospective studies. Castro et al. [4] address the extraction and classification of mammography information from radiology reports written in English. These free-text reports contain information expressed using the standardized lexicon and classification system described in BI-RADS (Breast Imaging Reporting and Data System). As described below in more details this study tackled the problem from an application point of view and attempted to cover a large number of hospitals.

Table 1 Best paper selection of articles for the IMIA Yearbook of Medical Informatics 2018 in the section 'Natural Language Processing'. The articles are listed in alphabetical order of the first author's surname.

Section Natural Language Processing
▪ Castro SM, Tseytlin E, Medvedeva O, Mitchell K, Visweswaran S, Bekhuis T, Jacobson RS. Automated annotation and classification of BI-RADS assessment from radiology reports. J Biomed Inform 2017 May;69:177-87.
▪ Pérez A, Weegar R, Casillas A, Gojenola K, Oronoz M, Dalianis H. Semi-supervised medical entity recognition: A study on Spanish and Swedish clinical corpora. J Biomed Inform 2017 Jul;71:16-30.
▪ Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What Patients Can Tell Us: Topic Analysis for Social Media on Breast Cancer. JMIR Med Inform 2017 Jul 31;5(3):e23.

Rare phenotypes and comorbidities often require specific work to be retrieved, because they are often undocumented or insufficiently documented in structured data. Two of the best paper candidates this year investigate methods to address this issue. Bejan et al. [18] present a large-scale study analyzing free text EHR content to extract a rare phenotype (homelessness and adverse childhood experiences) that is unlikely to be documented in structured data. The authors use query expansion based on word embeddings. The method has the potential to be applicable to other rare phenotypes. The study is notable for its large-scale use of clinical notes to study social determinants of health although the results of this particular case study are overall unsurprising. Boytcheva et al. [19] present a comorbidity mining approach for extracting co-occurring patterns at the patient level in a large collection of Bulgarian outpatient records. The method relies on features derived from both structured information and free text. Features from free text (including drug names and numerical values for body mass index (BMI), weight, and blood pressure) are extracted using regular expressions. The system is able to detect various interesting comorbidities and relations, some of which are validated in the literature.

Finally, one of the best papers addressed named entity recognition in Spanish and Swedish (Perez et al. described in more details below [2]). While an increasing number of tools and NLP pipelines become freely available for English, similar state-of-the-art tools are still needed for languages other than English and important research towards this goal includes methodological studies such as that of Pérez et al. who offer an in-depth study of supervised and unsupervised methods.

Papers with an Emphasis on Methods

A number of papers specifically explore NLP methods such as negation detection or corpus annotation for development. We outline a few here.

Detecting negation and its scope is an important problem in NLP [20]. It continues to motivate studies, including in languages other than English. Kang et al. [21] built a corpus of Chinese admission and discharge summaries annotated with negation and its scope. F-measure for negation detection was above 0.98 with a Conditional Random Field (CRF) supervised classifier. They tested the contribution to negation detection of word segmentation instead of simple character segmentation, and of embeddings instead of one-hot representations: word segmentation always helped, and embeddings helped when associated with word segmentation, reaching 0.99 F. Scope detection obtains 0.95 F, also based on a CRF, on negations found by the system. Garcelon et al. [22] used cue words and regular expressions to detect negated and family history sentences in French patient records. They ran the resulting detector on 1.6 million records to collect cases with "Lupus and diarrhea", "Crohn's and diabetes", or "NPHP1". Using negation and family history filters reduced the rate of false positives from 71% to 15%. This study shows the impact that negation and family history can have on simple case counts.

Supervised learning requires annotated corpora for system training, hence the importance of creating and distributing such resources. One of the candidate best papers, Alvaro et al. [23] created and made available a comparable pharmacovigilance corpus with two sources: Twitter and PubMed. By using the same annotation guidelines and expert annotators, they produced a homogeneous corpus that can be used to compare the expression of drugs, diseases, and symptoms in these two sources, and information extraction systems for pharmacovigilance. This work is timely given the large number of reported studies on social media and their potential interest for pharmacovigilance.

Concluding Remarks

One strong trend in 2017 publications was the revisiting of classic problems with deep learning and neural methods as for example, classification of mental health social media posts [24] or negation in Chinese [21]. Resource development remains crucial for applying these methods and for facilitating adaptation to new text genres, domains, or languages. Several studies produced annotated corpora as their main goal [23] or as part of their evaluation process. Generalization and transferability to other departments and institutions remains difficult, hence the interest for cross-institution studies [4,16,17].

References

In the reference list below, papers that were shortlisted as best paper candidates are marked with a *.

1. Filannino M, Uzuner Ö. Advancing the State of the Art in Clinical NLP through Shared Tasks. Yearb Med Inform 2018:184-92.
2. * Pérez A, Weegar R, Casillas A, Gojenola K, Oronoz M, Dalianis H. Semi-supervised medical entity recognition: A study on Spanish and Swedish clinical corpora. J Biomed Inform 2017 Jul;71:16-30.
3. * Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What Patients Can Tell Us: Topic Analysis for Social Media on Breast Cancer. JMIR Med Inform 2017 Jul 31;5(3):e23.
4. * Castro SM, Tseytlin E, Medvedeva O, Mitchell K, Visweswaran S, Bekhuis T, et al. Automated annotation and classification of BI-RADS assessment from radiology reports. J Biomed Inform 2017 May;69:177-187.
5. Norman C, Leeflang M, Zweigenbaum P, Névéol A. Automating Document Discovery in the Systematic Review Process: How to Use Chaff to Extract Wheat. Language and Resource Evaluation Conference, LREC 2018.
6. Gonzalez-Hernandez G, Sarker A, O'Connor K, Savova G. Capturing the Patient's Perspective: a Review of Advances in Natural Language Processing of Health-Related Text. Yearb Med Inform 2017:214-27.
7. * Cheng Q, Li TM, Kwok CL, Zhu T, Yip PS. Assessing Suicide Risk and Emotional Distress in Chinese Social Media: A Text Mining and Machine Learning Study. J Med Internet Res 2017 Jul 10;19(7):e243.
8. * Miller M, Banerjee T, Muppalla R, Romine W, Sheth A. What Are People Tweeting About Zika? An Exploratory Study Concerning Its Symptoms, Treatment, Transmission, and Prevention. JMIR Public Health Surveill 2017 Jun 19;3(2):e38.
9. * Lu Y, Wu Y, Liu J, Li J, Zhang P. Understanding Health Care Social Media Use From Different Stakeholder Perspectives: A Content Analysis of an Online Health Community. J Med Internet Res 2017 Apr 7;19(4):e109.
10. * Hao H, Zhang K, Wang W, Gao G. A tale of two countries: International comparison of online doctor reviews between China and the United States. Int J Med Inform 2017 Mar;99:37-44.
11. * Amith M, Cunningham R, Savas LS, Boom J, Schvaneveldt R, Tao C, et al. Using Pathfinder networks to discover alignment between expert and consumer conceptual knowledge from

online vaccine content. J Biomed Inform 2017 Oct;74:33-45.
12. Prud'hommeaux E, van Santen J, Gliner D. Vector space models for evaluating semantic fluency in autism. Proc ACL 2017:32-7.
13. * Kang T, Zhang S, Tang Y, Hruby GW, Rusanov A, Elhadad N, Weng C. EliIE: An open-source information extraction system for clinical trial eligibility criteria. J Am Med Inform Assoc 2017 Nov 1;24(6):1062-71.
14. * Iqbal E, Mallah R, Rhodes D, Wu H, Romero A, Chang N, et al. ADEPt, a semantically-enriched pipeline for extracting adverse drug events from free-text electronic health records. PLoS One 2017 Nov 9;12(11):e0187121.
15. Soysal E, Wang J, Jiang M, Wu Y, Pakhomov S, Liu H, Xu H. CLAMP - a toolkit for efficiently building customized clinical natural language processing pipelines. J Am Med Inform Assoc 2017 Nov 24.
16. * Ye Y, Wagner MM, Cooper GF, Ferraro JP, Su H, Gesteland PH, et al. A study of the transferability of influenza case detection systems between two large healthcare systems. PLoS One 2017 Apr 5;12(4):e0174970.
17. * Sohn S, Wang Y, Wi CI, Krusemark EA, Ryu E, Ali MH, et al. Clinical documentation variations and NLP system portability: a case study in asthma birth cohorts across institutions. J Am Med Inform Assoc 2017 Nov 30.
18. * Bejan CA, Angiolillo J, Conway D, Nash R, Shirey-Rice JK, Lipworth L, et al. Mining 100 million notes to find homelessness and adverse childhood experiences: 2 case studies of rare and severe social determinants of health in electronic health records. J Am Med Inform Assoc 2018 Jan 1;25(1):61-71.
19. * Boytcheva S, Angelova G, Angelov Z, Tcharaktchiev D. Mining comorbidity patterns using retrospective analysis of big collection of outpatient records. Health Inf Sci Syst 2017 Sep 28;5(1):3.
20. Wu S, Miller T, Masanz J, Coarr M, Halgrim S, Carrell D, et al. Negation's not solved: generalizability versus optimizability in clinical natural language processing. PLoS One 2014 Nov 13;9(11):e112774.
21. Kang T, Zhang S, Xu N, Wen D, Zhang X, Lei J. Detecting negation and scope in Chinese clinical notes using character and word embedding. Comput Methods Programs Biomed 2017 Mar;140:53-9.
22. Garcelon N, Neuraz A, Benoit V, Salomon R, Burgun A. Improving a full-text search engine: the importance of negation detection and family history context to identify cases in a biomedical data warehouse. J Am Med Inform Assoc 2017 May 1;24(3):607-13.
23. * Alvaro N, Miyao Y, Collier N. TwiMed: Twitter and PubMed Comparable Corpus of Drugs, Diseases, Symptoms, and Their Relations. JMIR Public Health Surveill 2017 May 3;3(2):e24.
24. Gkotsis G, Oellrich A, Velupillai S, Liakata M, Hubbard TJ, Dobson RJ, et al. Characterisation of mental health conditions in social media using Informed Deep Learning. Sci Rep 2017 Mar 22;7:45141. Erratum in: Sci Rep 2017 May 16;7:46813.

Correspondence to:
A. Névéol, P. Zweigenbaum
LIMSI
Campus universitaire, bât 508
Rue John von Neumann
F-91405 Orsay cedex
France
E-mail: {neveol,pz}@limsi.fr

Appendix: Content Summaries of Best Papers for the 2018 IMIA Yearbook, Section Clinical Natural Language Processing

Castro SM, Tseytlin E, Medvedeva O, Mitchell K, Visweswaran S, Bekhuis T, Jacobson RS

Automated annotation and classification of BI-RADS assessment from radiology reports

J Biomed Inform 2017 May;69:177-87

This paper presents a system for automatically extracting and classifying mammography information from radiology reports written in English. This is a well conducted study comparing rule-based and machine learning methods for BI-RADS (Breast Imaging Reporting and Data System) categories extraction from breast radiology reports (Conditional Random Field, F=0.95), together with their laterality (Partial decision trees, F=0.91-0.93). It can be noted that the study addresses types of clinical texts and entities that are less challenging than others. However, the corpus offers high text variety with reports from 18 hospitals. While the authors are neither experts on rule-based or machine learning methods, they present an excellent report of their work from the application point of view: describing the caveats of reproducing or adapting previous work, and using toolkits at their disposal towards the targeted application goal.

Pérez A, Weegar R, Casillas A, Gojenola K, Oronoz M, Dalianis H

Semi-supervised medical entity recognition: A study on Spanish and Swedish clinical corpora

J Biomed Inform 2017 Jul;71:16-30

This paper addresses named entity extraction from clinical corpora in Swedish and Spanish. It studies the influence of two types of unsupervised word representations on clinical information extraction performance: word embeddings as obtained by the word2vec method, clustered using K-means, and Brown clusters. The authors go beyond reporting experiments on two languages other than English and also offer methodological insight on the contribution of different classifiers and unsupervised features when little training data is available. The study is original in comparing the same set of configurations based on ensembles of word representations in both corpora, although different types of entities are annotated in the two languages (Drug/Diseases for Spanish, and Body Part/Disorder/Finding for Swedish).

Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T

What Patients Can Tell Us: Topic Analysis for Social Media on Breast Cancer

JMIR Med Inform 2017 Jul 31;5(3):e23

This study reports on social media analysis relying on two corpora on the topic of breast cancer in French. This paper presents a strong use case and good explanations of why the work is significant from multiple points of view: social media, quality of life in cancer patients, and clinical questionnaire development. The authors use Latent Dirichlet Analysis (LDA) to detect the topics in breast cancer messages from Facebook groups and in forum posts. After a balanced analysis of the LDA results, the automatically identified topics are aligned with internationally standardized self-administered questionnaires used in cancer clinical trials in order to validate the results and identify gaps in the questionnaires. The study draws conclusions that may bring an impact on the maintenance of the international questionnaires.