e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-58

Towards a decision support system for hand dermatology

Luca MAZZOLA[a,b1], Alice CAVAZZINA[a], Francesco PINCIROLI[a], Stefano BONACINA[a], Paolo PIGATTO [c], Fabio AYALA [d], Ornella DE PITÀ [e], and Sara MARCEGLIA[a,2]

[a] *eHealthLAB, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, ITALY*
[b] *Università della Svizzera italiana - USI - Via Buffi, 13, 6900 Lugano (TI) Switzerland*
[c] *Università degli Studi di Milano, Milano, Italy*
[d] *Università di Napoli Federico II, Napoli, Italy*
[e] *Servizio di Allergologia ed Immunologia Clinica e di Laboratorio, Istituto Dermopatico dell'Immacolata, Roma, Italy*

Abstract. The complexity of the medical diagnosis is faced by practitioners relying mainly on their experiences. This can be acquired during daily practices and on-the-job training. Given the complexity and extensiveness of the subject, supporting tools that include knowledge extracted by highly specialized practitioners can be valuable. In the present work, a Decision Support System (DSS) for hand dermatology was developed based on data coming from a Visit Report Form (VRF). Using a Bayesian approach and factors significance difference over the population average for the case, we demonstrated the potentiality of creating an enhanced VRF that include a diagnoses distribution probability based on the DSS rules applied for the specific patient situation.

Keywords. Decision Support System (DSS), hand dermatology, differential diagnosis, diagnostic implicit knowledge, medical practices.

Introduction

In dermatology, hand skin pathologies often present similar signs and symptoms [1] that can be interpreted by highly experienced and trained specialists. Hence, inexperienced general practitioners (GPs) or non-specialized dermatologists need support in performing their diagnosis, and in timely recognizing the disease to refer the patient to an appropriate specialized centre. In fact, wrong diagnoses and thus inadequate therapies or treatments can worsen patients' conditions. Recently, the differential diagnostic process ([2], [3], [4]) has been applied to hand skin pathologies,

1 Corresponding Author. Email: mazzola.luca@gmail.com – luca.mazzola@polimi.it
2 Email: sara.marceglia@polimi.it On behalf of the "Ambulatorio della Mano": S. Amato, Ospedale Civico di Palermo; G. Bornacina, Azienda Ospedaliero- Universitaria Maggiore della Carit´a, Novara; P. Calzavara-Pinton, Spedali Civili di Brescia; S. Calvieri, La Sapienza Universit´a, Roma; M. Congedo, Ospedale Vito Fazzi, Lecce; A. Cristaudo, Istituto Dermatologico San Gallicano IRCCS, Roma; C. Crosti, Universit´a degli Studi di Milano; F. Cusano, Ospedale Rummo, Benevento; M. Gola, Universit´a degli Studi di Firenze; P. Lisi, Universit´a degli Studi di Perugia; P. Patrone, Azienda Ospedaliero-Universitaria, S. Maria della Misericordia, Udine; C. Potenza, Polo Pontino, La Sapienza Universit´a, Roma; M. Travaglini, Ospedale A. Perrino, Brindisi.

but a systematic assessment in a multi-centre trial has not been performed yet. For this reason, the multi-centre study e-Dermatology - involving 16 Italian centres of excellence in the area of hand dermatology - was designed aimed at collecting information (enriched with implicit knowledge) regarding the practices of differential diagnosis by highly specialized dermatologists. The information collected in e-Dermatology can be used to develop a decision support system (DSS) for the hand dermatitis differential diagnosis devoted to non-specialized dermatologists and GPs, providing them with a simple tool for a first supported diagnosis of these diseases.

In this work, we present the e-Dermatology system for the collection of clinical diagnostic data in patients with hand skin pathologies and its use to define a preliminary set of decisional rules that can be used in a DSS aimed to support GPs and non-specialized dermatologists in their diagnostic process.

1. Methods

Sixteen Italian clinical centres specialized in hand skin pathologies participated in e-Dermatology, providing, in one year, data from 900 patients, in 1300 visits. The pathologies considered are divided in two classes: primary hand skin pathologies (Allergic Contact Dermatitis, ACD, Irritant Contact Dermatitis, ICD, ICD and ACD, Atopic eczema. Atopic eczema and ICD, Vescicular eczema, Hyperkeratosis eczema, Nummular eczema) and secondary hand skin pathologies in which the hand signs and symptoms are the epiphenomenon of another disease (Psoriasis, Palmoplantar keratoderma, Superficial cutaneous mycoses, Continuous acrodermatitis, Pityriasis rubra pilaris). As a first step, we designed and developed a specialized Visit Report Form (VRF) system [5], to support data collection. Then, we analysed the information collected through the shared VRF to disclose some possible rules for decision support.

VRF implementation: The VRF contents were designed together with trained specialists [6] belonging to the Italian reference centres involved in the study, to ensure its clinical correctness and usefulness. Then, the implemented VRF was re-assessed by specialists in the field who did not participate in its design, to prevent from possible biases introduced by personal experiences of the designers. The VRF comprised an explicit tracking system that recorded a personal score(in a 0-5 range) given to the collected information by the specialist during a patient encounter. The question underlying the scoring phase was: how much was this information (e.g. patient's gender, age, parts of the hand with altered skin) important/useful to diagnose patient's disease?

The web-based VRF was developed to ensure patient's anonymity and homogeneous data collection.

Data Analysis and rule definition: Data collected through the developed VRF system, after cleaning and scrubbing operations were analysed to obtain possible rules to suggest the diagnosis given an information dataset, and also the next possible diagnostic step (e.g., a test to be prescribed to reach a diagnosis, or a symptom to monitor).

Each field of the VRF was considered as a variable of the diagnostic process. The variables could have a binary value (e.g., yes/no, male/female), a categorical value (multi-choice field), or a numeric value (e.g., clinical scales). The distribution of each variable with respect to the diagnosed pathology was calculated in order to identify the most significant variables in the diagnostic process. Also, we evaluated the distribution

of the diagnoses among the combination of some variables selected following some clinical considerations.

Then, we calculated the conditional probability of a certain diagnosis provided a certain variable or a certain variable combination. To do so, we relied on a well-known statistical approach to estimate non-observable relationships and variables through the observation of depending events realization, called inferential statistic [7]. The output of this process is a Bayesian Network, an acyclic directed graph that encodes cumulated probabilities and that can be used to estimate the probability of depending events knowing the probabilities of the hidden factors [8].

Based on the results from the previous step, we calculated the conditional probabilities of being diagnosed of certain pathology, provided a certain attribute set. We then multiplied these probabilities obtained for each variable/combination of variables, and obtained a general probability for a variable set, that represent a binary flag of the possibility to observe the specific disease in the current case. Through this approach we found the diagnoses that should be excluded (total probability = 0). Then, to obtain a better discrimination amongst the non-excluded pathologies, we considered the difference between the conditional probability of single variables –or variable sets, based on medical considerations– within the current case and the average probability for the diagnosis in the population. In this way, we could estimate whether the current value of the single variable/variable set enhanced or not the population probability for a certain diagnosis. We used this information to apply a weighting factor that left untouched the variables supporting the diagnosis and weighted negatively the adverse ones. This means multiplying each factor for a discount value of +1 if the conditional probability was higher or equal than the population probability, and of -1 otherwise).

Eventually, by the combination of the conditional probabilities and a weighting process based on the sample size and on the significance of the variable in respect to its discriminatory capability, we created a possible set of rules able to suggest the most likely diagnosis, sorted by decreasing confidence.

2. Results

The Visit Report Form System: The VRF system was divided in 5 different sections: personal information, demographical information, physiological conditions, pathological conditions, and contact reaction tests [optional]. The VRF (Fig.1, *left side*) also included clinical scales such as the modified Total Lesion Symptom Score (mTLSS). A new VRF was created once a patient was visited in one of the clinical centres. Hence, each patient could have more than one VRF, each corresponding to a visit. The final diagnosis could be reached after one or more visits.

After the access by protected protocol to the website using the authentication codes provided, each centre could insert a new visit, retrieve the ones already executed, or insert an update for an already enrolled patient. To guarantee patient's privacy, the identification of the patient was achieved through an ID assigned by the centre, and by managing the association of the ID with the patient's name outside the system. Also, the system, to guarantee a secure identification of the patient, required inserting the patients code twice for crosschecking. In the case the patient code already existed, the persistent information (such as birth date, residence, gender) were preloaded in the VRF, making the filling out of new visit report form quicker.

Results and rules: Of the 1300 VRFs collected, 85% was error-free and considered for the subsequent analyses. In the other 15% either storage errors or errors during VRF filling-in occurred, and corresponding VRFs were excluded. The distribution of diagnoses was not homogenous: 38% of the population was affected by ACD, 20% by ICD, 19% by eczema, 20% by psoriasis, whereas the other pathologies covered the remaining 3%.

We considered as main diagnostic variables gender, contact with animals/pets, patient's job, family history, affected sites (palm, back, only hand, etc) and distribution of lesions, patient's reported factors that could have caused the pathology (e.g., contact with soaps other irritative substances), the mTLSS scale, and the results of the histology examination (if present). Their distribution among diagnoses showed that some factors were strongly correlated with some pathologies.

For instance, the superficial skin mycosis was the disease that had the highest percentage of positivity in the mycological test (70% of the population affected present a positive test result); the likelihood of developing the ACD in the case of family history of palmoplantar keratoderma, was quite high, summing up to 82% of the analysed cases; individuals having the combination of lesions on palms (site of infection at the time of visit) - erythema, desquamation (type of injury) - confluent in rashes (location to the injury) were affected by psoriasis with the rule confidence of almost 100%.

By applying then the weighting strategy presented in the previous section, we obtained a ranking of the probability associated to each pathology provided certain variable or combination of variables values for the specific patient. The validation of the system was then performed through the classification of some medical records extracted randomly from the same dataset (coming from the e-Dermatology multicentre study). The system gave satisfactory results as the first diagnosis was correctly indicated in the 70% of cases and, for the remaining 30%, the actual (real) diagnosis was ranked inside the possible ones, even if not as the most probable (primary).

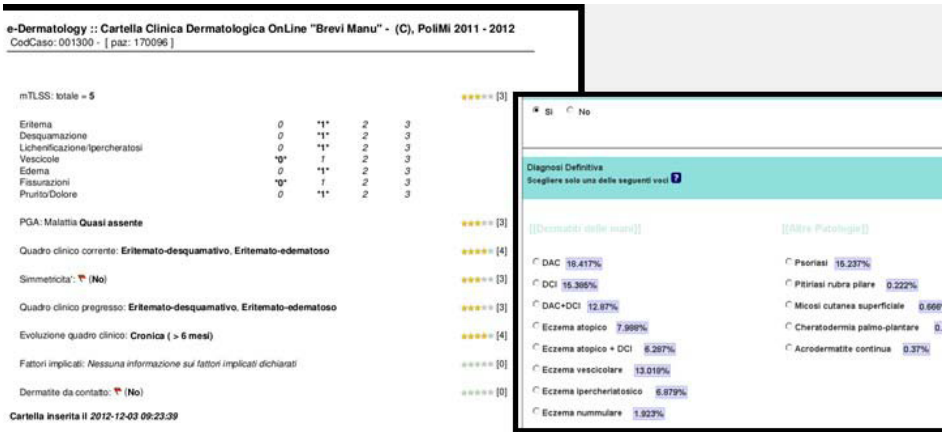

Figure 1. *Left side*: Representation of the mTLSS and PGA scales inside a compiled VRF. *Right side*: A possible VRF enrichment by the DSS rules: probability suggestion for each pathology.

Application to the VRF: The results obtained can be integrated in the implemented VRF (Fig.1, *right side*). The probability of each diagnosis, expressed as a percentage,

can be shown and updated each time the clinician fills-in one of the VRF fields. The probabilities are not presented as ranked by their order, to enforce the idea that the real choice is performed by the – human – medical professional. In fact, the percentages are indicative values from the data extracted, based on statistical and computational methods [7], in accordance with the clinical experience of the experts involved in the study.

3. Discussion

Our results provide the basis for a system able to support the GP or the non-specialized dermatologist in diagnosing hand-skin pathologies. This additional information can support the clinicians by allowing them to better target their diagnosis or to suggest a visit with a specialist, in case of doubt. The system is however still a prototype. In fact, despite being based on a large number of cases, the sample population was still biased by the incidence of the pathologies examined. The rules developed can be hence further optimized increasing the number of cases representing rare pathologies. Moreover, the proposed supportive system can be enhanced by providing the diagnosis probabilities while the clinician fills-in the VRF, instead of providing it at the end.

In conclusion, the significance of this work is twofold: first, this could help to increase the correctness rate of diagnosis performed by the GPs - or at least to make them aware of the need of a second specialized consultation. This decreases the risk of suggesting a wrong recovery path that could negatively impact on disease progression. Second, the collected information and the implicit diagnostic pathway (or rules) can lead to a better understanding of the usual way of reasoning followed by specialists for hand skin disease diagnosis.

References

[1] A. Svensson, M. Lindberg, B. Meding, et al, Self-reported hand eczema: symptom-based reports do not increase the validity of diagnosis. *Br J Dermatol*, 147 (2002): 281284.

[2] Weisshaar, E., Kallen, U., Weiss, M.: 'The itching hand"- important differential diagnoses and treatment. *J. Dtsch. Dermatol. Ges*. 11(1), 31–42 (2013), doi:10.1111/j.1610-0387.2012.08002.x.

[3] Jackson, S., Nesbitt, L.T.: Differential Diagnosis for the Dermatologist. Springer (2012) ISBN: 9783642280061

[4] Pesut, D.J., Herman, J.: Metacognitive Skills in Diagnostic Reasoning: Making the Implicit Explicit. *Int J Nurs Terminol and Classif* 3, 148–154 (1992), doi:10.1111/j.1744-618X.1992.tb00530.x

[5] L. Mazzola, S. Marceglia, S. Bonacina, et al 'Explicit tracking in the diagnostic process for hand dermatological practices'. In Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management. Healthcare and Safety of the Environment and Transport (2013), Springer Berlin Heidelberg : pp. 248-257.

[6] F. Ayala, M. Nino, G. Fabbrocini, L. et al.: Quality of life and contact dermatitis: a disease-specific questionnaire. *Dermatitis* 21(2), 84–90 (2010); PubMed PMID: 20233546

[7] Cox D.R., Principles of Statistical Inferences. CUP (2006). ISBN: 0-521-68567-2

[8] Pearl J., Causality: Models, Reasoning, and Inference. Cambridge University Press (2000). ISBN: 0-521-77362-8

[9] D. Eynard, L. Mazzola, A. Dattolo, Exploiting tag similarities to discover synonyms and homonyms in folksonomies. In: Software: Practice and Experience. John Wiley & Sons, Ltd. (2012), doi: 10.1002/spe.2150