RELEVANCE OF LYMPHOSCINTIGRAPHY QUANTIFICATION IN COMBINED DIAGNOSTICS OF UPPER LIMB LYMPHEDEMA

Petr Gebouský

[1]*Department of Adaptive Systems*
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Prague, CZECH REPUBLIC
E-mail: gebousky@utia.cas.cz

Abstract: Secondary upper limb lymphedema, a frequent complication after breast cancer therapy, can be successfully treated only when it is diagnosed in its early stage. Use of, otherwise well-established, lymphoscintigraphically supported staging is inhibited by a slow lymphatic dynamics of upper limbs, which allows a routine collection at most three images reflecting it. The proposed Bayesian staging methodology, described in the paper, relies on a simplified accumulation model to get quantitative lymphoscintigraphy and uses normal probabilistic mixtures for a computerized disease staging that exploits fully the routinely available information.

Keywords: quantitative lymphoscintigraphy, secondary upper limb lymphedema, Bayesian evaluation

1 INTRODUCTION

Lymphedema, the edema caused by lymphatic system insufficiency, is a chronic disease that is frequently misdiagnosed, treated too late, or not treated at all. At the same time, the success and efficacy of its therapy depend strongly on the disease stage. The upper limb lymphedema often arises as a secondary complication of the breast cancer therapy and axillary lymph node dissection. The treatment of the breast cancer concerns of a significant group of patients, for instance, yearly about 4000 women considering Czech Republic only. Frequency of the lymphedema incidence is relatively high, about 5%-30% (Szuba *et al.*, 2003). This state calls for an efficient and reliable diagnostic method allowing a safe recognition of early lymphedema stages.

Besides the basic clinical assessment, lymphoscintigraphy seems to be the adequate sensitive inspection method already used for judging the state of lymphatic system. The qualitative evaluation characterizes well lymphatic morphology. On the other hand, the recognition of treatment-critical latent disease stages is rather difficult (Weissleder and Weissleder, 1988). An increase of sensitivity and diagnosis accuracy is commonly expected from a quantitative evaluation. Yet, this expectation has been fulfilled only partially. Slow dynamics of the upper-limb lymphatic system makes its diagnostics specific. It reduces severely the number of measurements available for dynamic study that can be routinely taken. The limitation is caused by, always restricted, time-capacity of the gamma camera as well as by the limited ability of patients to undergo a sufficiently rich series of inspections within the time interval covering dynamics of the upper-limb lymphatic system. Consequently, two or three images taken on each limb is the realistic, routinely accessible, number of images.

The small amount of data available makes a diagnostic inference hard. It also makes the evaluation of traditional physiological indicators very volatile. Consequently, no reliable, clinically

accepted, quantitative evaluation of the upper-limb lymphoscintigraphy is at disposal. The only known way to counteract the lack of measured data is a careful modeling and data processing. Inspired by the depot-clearance rate technique (Pain *et al.*, 2002) a new method of quantification has been proposed in (Gebouský *et al.*, 2003). The depot clearance method models the dynamics of the colloid accumulation at the injection site only. The discussed approach is based on modeling of the colloid accumulation within the remaining parts of the limb. The adopted inspection of the various limb parts respects the evidence that lymphedema may appear locally within the limb.

Considering the small amount of uncertain data, the use of Bayesian methodology (Berger, 1985) which substitutes lack of data by prior knowledge, appears to be the only viable possibility. The necessary simplified modeling of the accumulation of the radiotracer within the limbs drifted by lymphatic flow was done. Processing of the model within the Bayesian decision-making paradigm resulted in the routinely applicable quantitative lymphoscintigraphy of the upper limb lymphedema.

Efficiency of the proposed quantification method has already been confirmed (Gebouský, 2003). The results indicate that the method increases the diagnosis accuracy. However, the reliable detection of early disease stages cannot be reached without employing all routinely available, clinical and scintigraphic information. A combination of these disparate information sources and judging performance of the resulting procedure form the core of this paper. Specifically, the paper describes the way of an automatic combination of these disparate data and verifies the impact of the proposed quantification methodology on the decision about the lymphedema staging. The discussed combination of such data is based here on an exploitation of probabilistic mixtures (Kárný *et al.*, 2005) to this purpose.

2 DATA COLLECTED

Each patient underwent both the lymphoscintigraphic and clinical evaluations. The scintigraphic data were processed by the referred quantitative methodology (Gebouský, 2003) and supplemented by the information characterizing the previous treatment and therapy. The evaluation extends the qualitative inspection of scintigraphic data, which is already a decisive part of the routine inspection.

2.1 Clinical Assessment

The data concerning clinical evaluation include categorized therapy history, subjective feelings of the patient and clinical findings.

2.2 Lymphoscintigraphy

Lymphoscintigraphic inspections were performed for all patients on both upper limbs.

Imaging and Data In all subjects 20 MBq of 99mTc-labeled colloid in the volume 0.1-0.2 ml was injected subcutaneously into the first and fourth inter-digital web space of each hand. Then the initial calibrating 60-seconds image of the injection site is acquired. The remaining three images are routinely obtained within the range 30-180 minutes after administration. The whole arm is imaged in a supine position. Images of both upper limbs are collected for 60 seconds by LFOV gamma camera.

For the quantitative evaluation, the regions of interest (ROI) are drawn around the axillary and supraclaviculary region, the forearm and upper arm. The accumulated activities within ROIs are characterized by the total integral counts over them. These counts are corrected to the physical decay of the traces over the inspection period.

Qualitative Evaluation The trained nuclear-medicine expert bases the qualitative evaluation of lymphoscintigraphy on the visual assessment of the images. The expert tries to recognize the following patterns: the number of visible arm and cubit nodules; the lack of a transit in the application site; the visibility of the extended lymphatic vena and the existence of dermal backflow. The level of the dermal backflow and its local position are differentiated. The outlined qualitative comparison of scintigraphic images leads to the expert's assessment S of the lymphedema stage.

3 QUANTIFICATION METHOD

3.1 Radiotracer Accumulation Model

The proposed quantification technique relies on the simplified modeling. It respects the small amount of available data reflecting the accumulation within the predefined ROIs covering forearm, upper arm and axilla.

The relative scintigraphic responses within the individual ROIs are described by discrete-time input-output dynamic model. The relative scintigraphic response is the time curve of the activity of the accumulated colloid normalized by the administered activity when the modeled system is stimulated by unit impulse. The response (x) forms time activity curve (TAC). Specifically, a cascade of first-order linear models, with a common dynamical parameter a for each of the d sections, and with a common gain parameter b is chosen. It is a flexible compromise between the need to characterize the complex distributed nature of the lymphatic system and the need to get a model with a few unknown parameters. TAC x_t at time $t = 0, 1, \ldots$ is related to the model parameters by the formula:

$$x_t = b\binom{t+d-1}{t}a^t, \;\; t \geq 0. \tag{1}$$

While this expression models the whole response, the measurements are taken only in small subset of discrete time moments. Noisy samples (y_t) of the TAC are observed. The aggregation of data counts permits to characterize the overall noise effect by the additive zero-mean normal noise (e_t), i.e., $y_t = x_t + e_t$. The considered infrequent measurement implies that the noise samples can be treated as conditionally independent. The normalization implies that the noise variance r can be assumed approximately constant.

3.2 Prior Information on Model Parameters

The TAC (1) and the additive noise e_t are characterized by the unknown parameters $\Theta = (a, b, d, r)$. The noise variance r reflects the measurement process. Thus, for its estimation, data from various ROIs and patients can be used. The remaining three parameters are strictly the patient specific. They depend both on the modeled limb and ROI and have to be estimated using the available two or three measurements. This is obviously impossible without prior information. Its systematic use is the key advantage of the Bayesian paradigm adopted for the inference on diagnostically significant quantities from the sparse data.

The prior information is rich in the inspected problem and can be expressed through intervals of a priori possible values of Θ. Let us list and briefly comment prior information on the patient-specific parameters:

- d ($1 < d \leq 6$) - the parameter describes the penetration rate through the limb and modifies the shape of the TAC. For $d = 1$, the model would coincide with the exponential model used for depot clearance method. The chosen upper bound 6 is very conservative guess.

- a $(0 < a < 1)$ - the parameter determines model dynamics. The specified interval reflects the fact that the inspected responses are stable and non-oscillatory. In implementation, this interval is shrunk to reflect the observed slow accumulation dynamics. Typically $0.9 < a < 0.999$: the specific range depends on the order d.
- b $(0 < b < 1)$ - the parameter describes the model gain. The response is non-negative and cannot exceed the applied input. Taking into account that no activity is created within the limb, a tighter upper bound bmax can be evaluated as a function of the inspected a, d (Gebouský, 2003).

The above information items were converted into a complete prior distribution on the patient-specific unknown parameters (a, b, d). Since no further detailed information is available for them, uniform distributions on the above ranges were chosen. This choice can be justified via the principle of insufficient reasons (Jeffreys, 1985). Sufficient amount of data available for estimation of the common noise variance r has allowed us to select for its description computationally advantageous conjugate prior (Berger, 1985).

3.3 Processing of Information Sources

With the measured data, the chosen model and the prior distribution, a relatively straightforward Bayesian evaluation of posterior distributions (Berger, 1985) provides point estimates as well as their precisions. The patient-specific parameters as well as the noise-free time activity curve x_t at *any* discrete time moment are estimated on each individual ROI.

The reconstruction of the TAC motivated the outlined modeling and estimation. The success in this respect and practical needs turned our attention towards the disease staging, which is predominantly addressed in this paper. This is a classical difficult pattern-recognition problem (19). It requires selection of quantifiers (features) that allow reliable differentiating stages of lymphedema. The estimate of the triple of the patient-specific parameters (for each ROI) is a natural candidate to this purpose. Besides the estimates of time constant a, the gain b and the number of sections d shaping TAC, the value and position of the TAC maximum are considered as quantifiers. They represent the counterpart of the late uptake and appearance time used in (Weissleder and Weissleder, 1988). The residence time, widely accepted in nuclear medicine as a quantitative characteristic of accumulation kinetics is tested too. With the adopted scaling, the residence time in minutes is found as the area under the TAC estimate.

Note that in all cases just point estimates are passed to the further processing described bellow.

4 DISEASE STAGING AND RELEVANCE OF QUANTIFICATION

The computerized support of the staging assessment is the ultimate aim of the discussed data processing. An algorithm mapping the available patient data on a reliable estimate of the disease stage is searched for. A justified selection of the significant items within the data record D containing clinical, qualitative and quantitative scintigraphy results is the most important sub-problem addressed. Importance of items arising from the proposed quantitative lympho-scintigraphy is of a special interest.

The combination of clinical, qualitative and quantitative scintigraphy provided records D with 41 meaningful data items for each inspected limb. The subjective staging S, made by scintigraphic expert complements the record.

The constructed “staging” algorithm has to cope with sparse data. Even when we take data from different limbs attached to different subjective staging we get 176 records, each with 41 explanatory variables and single predicted variable. This makes us to use again the Bayesian

solution of this specific pattern recognition problem. Normal probabilistic mixtures are selected as the needed probabilistic model relating the unknown stage S to data D. The following reasons singled out this model class:

- Probabilistic models "naturally" model discrepancies of experts' opinion as well as always partially arbitrary boundaries between respective disease stages. Moreover, the modeled relationship is stochastic.
- The modeled staging concentrates around nominal discrete values $\{0, 1, \ldots, 4\}$ and the explanatory data contain continuous-valued entries. Thus, the hard problem of logit regression is faced, e.g. (Albert and Chib, 1993). Such a regression can be, however, approximated by a normal mixture whenever variances of the individual normal components are taken small enough.
- Normal mixtures formally coincide with artificial neural network made of radial basis functions. This network is known to approximate (almost) any multivariate mapping (Haykin, 1994). Thus, normal mixtures suit for the considered exploratory data analysis.
- Efficient algorithms exist for an approximate Bayesian estimation of normal mixtures, including estimate of its structure and model validation (Kárný *et al.*, 2005).

The processing was performed by Jobcontrol system (Tesař and Novák, 2005), which arisen from a software base Mixtools that covers all major tasks related to mixture estimation.

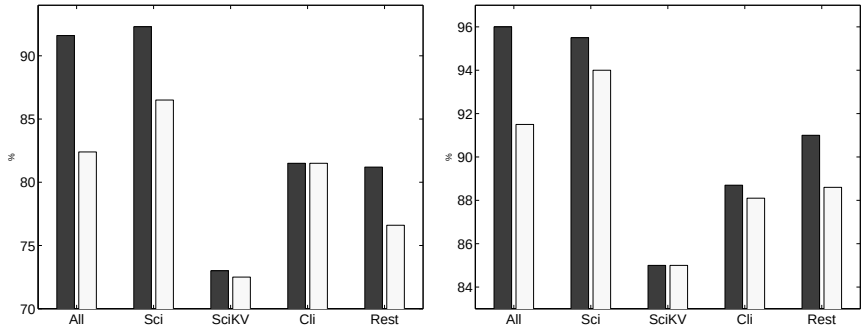
5 RESULTS & DISCUSSION

The evaluation results concern 88 patients, i.e., 176 limbs. The results are summarized here in Figure 1. The classification success is measured by the percentage of the coincidence with expert classification S.

Columns in figures correspond with a different extent of the exploitation of available explanatory data. *All* means exploitation of all of them. *Sci* uses combination of data provided by scintigraphic expert (except of his/her staging) and by the quantitative scintigraphy, i.e., by numerical characteristics of the estimated model (1). *SciKv* refers to the use of the later data only. *Cli* relies on data provided by clinician. *Rest* predicts the stage using the combination of data provided by the scintigraphic expert and data provided by clinician (again except of staging). Left columns correspond with results obtained when all considered data were used for estimating the mixtures, which model the relationship of explanatory data to the predicted lymphedema stage. Right columns reflect results when using the leave-one-out cross-validation.

Figure 1(a) displays correspondence of the classification by the estimated mixture with that provided by scintigraphic expert. Figure 1(b) gives up rarely populated higher stages of lymphedema and distinguishes just the statements: the limb does not have, 0, and the limb does have lymphedema, 1.

The results imply the following observations. Use of all scintigraphic explanatory data is the must. The evaluation of only clinical data or only data from quantitative scintigraphy is insufficient for reliable staging. The overall number of learning data is still too small as it seen on poorer cross-validation results. This also explains higher robustness of simpler models (*All* vs. *Sci*) or models dealing with discrete-valued data only (*Cli*). The quality of dichotomy evaluation is, however, very high.



(a) (b)

Fig. 1: Correspondence of (a) classified stage and (b) classified lymphedema 0/1.

6 CONCLUSIONS

The derived methodology of computerized lymphedema staging confirmed (among other outcomes) that: (i) Scintigraphic evaluation is necessary for low-stage cases. (ii) The characteristics of the model (1) are insufficient alone for staging but contributes significantly to qualitative scintigraphy in the way, which cannot be substituted by clinical evaluations only. (iii) Possibility of a finer computerized staging is indicated but richer data sets will be needed to convert the possibility into certainty. (iv) Finally, combined data from qualitative and quantitative scintigraphy can be used for a very reliable indication of lymphedema presence even in early stage.

ACKNOWLEDGEMENT

This research has been supported by IGA MZ ČR, contract number NC/7601 - 3, by GA ČR 102/03/0049 and by Research Center DAR, contract number MŠMT 1M0572. Thanks to Czech-Slovenian project MŠMT 8-2006-06 ”Data-driven modelling for decision making support and process monitoring” attendance on this workshop could be realized.

REFERENCES

Albert, J. and Chib, S. (1993), ‘Bayesian analysis of binary and polychotomous response data’, *Journal of the American Statistical Association* **88**(422), 669–679.

Berger, B. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.

Gebouský, P. (2003), Bayesian Quantitative Lymphoscintigraphy of Upper Limbs, PhD thesis, University of West Bohemia, Pilsen, Czech Republic.

Gebouský, P., Kárný, M. and Quinn, A. (2003), Lymphoscintigraphy of upper limbs: A Bayesian framework, *in* J. Bernardo, M. Bayarri, J. Berger, A. David, D. Heckerman and A. Smith, eds, ‘Bayesian Statistics 7’, Clarendon Press, Oxford, pp. 543–552. Proceedings of the Seventh Valencia International Meeting, June 2–6, 2002.

Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan, New York.

Jeffreys, H. (1985), *Theory of Probability*, Clarendon Press, Oxford.

Kárný, M., Bohm, J., Guy, T. V., Jirsa, L., Nagy, I., Nedoma, P. and Tesař, L. (2005), *Optimized Bayesian Dynamic Advising: Theory and Algorithms*, Springer, London.

Pain, S., Nicholas, R., Barber, R., Ballinger, J., Purushotham, A., Mortimer, P. and Peters, A. (2002), ‘Quantification of lymphatic function for investigation of lymphedema: Depot clearance and rate of appeareance of soluble macromolecules in blood’, *The Journal of Nuclear Medicine* **43**(3), 318–324.

Szuba, A., Shin, W., Strauss, H. and S.Rockson (2003), ‘The third circulation: Radionuclide lymphoscintigraphy in the evaluation of lymphedema’, *Journal of Nuclear Medicine* **44**, 43–57.

Tesař, L. and Novák, M. (2005), Support environment for system identification and controller design: Jobcontrol, Technical Report 2138, ÚTIA AV ČR, Praha.

Weissleder, H. and Weissleder, M. (1988), ‘Lymphedema: Evaluation of qualitative and quantitative lymphoscintigraphy in 238 patients’, *Radiology* **167**(3), 729–735.