From Pap Classification to Bethesda 2001 – Relating the Development of a Diagnostic Lexicon to Triage Systems

Amrita Mishra

Abstract

The Bethesda System (TBS) of cervical-vaginal nomenclature has been developed primarily to standardize descriptive terminology for reports on cytological examinations such as the Pap smear. This was done in part to prevent confusion in diagnosis and specimen adequacy evaluation. TBS has made a significant impact on laboratory practice. The popularity of the system among cytologists and clinicians signifies a thrust to establish certain protocols to standardize gynaecological patient management. The paper will delineate some aspects of the relationship between the TBS diagnostic lexicon, issues of risk management and triage for cervical-vaginal diagnosis, treatment and healthcare resource allocation, and the incremental consolidation of the networks involved.

Keywords:
Bethesda, TBS, cervical, cervix, cancer, cytology, triage, diagnosis, Pap, Papanicolaou, smear

Introduction

It is generally accepted that cytological screening for precancerous lesions of the cervix is effective in reducing incidence and mortality for cervical cancer. Such screening meets two criteria regarded as marks of a good screening procedure – relatively low cost and acceptance across populations, although the latter is significantly qualifiable in societies and groups where such cervical-vaginal monitoring is regarded as a breach of female modesty.

Cervical cancer is associated with a long lead-time in its inception and development. Precancerous lesions are seen as progressing through identifiable stages prior to invasive malignancy. The treatment options available for prevention of invasive disease are defined according to the stage detected. Scientific evidence for the efficacy of cervical cytology in

reducing the incidence of invasive disease and mortality comes from a wide range of non-experimental studies, specifically observational and case-control studies.

From the time of the introduction of the Pap smear to the present, the diagnostic terminology used to describe the cellular changes observed has been an area of contention. G. N. Papanicolaou described the earliest terminology, introduced in 1954 (Papanicolaou 1954, 20–21).

> (…) Class I – Absence of atypical or abnormal cells. Class II – Atypical cytology but no evidence of malignancy. Class III – Cytology suggestive of, but not conclusive for, malignancy. Class IV – Cytology strongly suggestive of malignancy. Class V – Cytology conclusive for malignancy.[1]

Since the 1950s, many different terms in different classifying schemata have been used to describe the contents of the cervical samples collected, slide-mounted and examined microscopically. There was much criticism of the ensuing 'Rashomon Effect', that the muddle of classification systems and languages created a conflict over the readings of tissue specimens, and thereby also on the modalities of sorting out patient prognoses and organizing treatment (Clarke & Casper 1996).[2]

In December 1988, a group of cytopathologists, pathologists, clinicians, and representatives of professional organizations related to gynaecological treatment and diagnosis attended a workshop sponsored by the National Cancer Institute at Bethesda, Maryland, USA. The purpose of the workshop was to address the 'diagnostic chaos' – the Rashomon effect – in cervical-vaginal cytology. This was to be done by recommending a *system of uniform descriptive diagnostic terminology*. The result of the workshop was therefore not just a listing of diagnostic terms, but a comprehensive modality of reportage and action for cervical-vaginal cytology. Criteria were set out for both evaluation of specimen adequacy and a reliable descriptive diagnosis – the whole underpinned by a driving concern with standardized actions, with defined clinical relevance, across the sites for the application of this system. This system became known as the Bethesda System (TBS) of cervical-vaginal nomenclature. TBS was introduced to quell the polysemic use of diverse diagnostic terms, what many specialists think of as diagnostic chaos.

The growing popularity of the system among cytologists and clinicians signifies the medical community's thrust to establish a uniform basis for patient management. TBS was introduced in 1988, and incrementally revised in 1991 and again in 2001.[3] The revised TBS was developed through a committee-based process of literature review, solicitation of expert opinions, and discussion of revisions on an e-forum.[4] TBS 2001 included revisions in statements of specimen adequacy, general categorization, and interpretation and results of epithelial cell abnormalities.

TBS establishes the *operational protocol* that makes Pap a tool in cancer screening. It enables three things. *One*, a specimen that meets conditions of adequacy is to be obtained from the patient and submitted to the laboratory.[5] *Second*, the specimen is to be properly prepared, screened, and interpreted. *Third*, the cytologists communicating their findings to the clinician must do so in a shared technical parlance; minimal or negligible room for manoeuvre with terminology. The framework for communication must reduce polysemy of technical terms to a bare minimum. This is related to the making of shared epistemic spaces, collaborations between cytologists and clinicians, and regulating information flows between different professional camps. This has never been an easy task, given the prevalence of rivalries and acrimony between such camps. Clarke and Casper discuss the opposition to Papanicolaou's exfoliative cytology by traditional pathologists (Casper & Clarke 1998, 259).[6]

In order to achieve a commonality of access to information across professional communities, the differences of morphology of normal cells and lesions have to be communicated within and between work settings that must therefore share both language and perceptions. The terms used to describe morphology would have to enforce an avoidance of subtleties. This could be achieved by setting out a clear-cut lexicon of terms, marked by a certain paucity of vocabulary. The semantics cannot be allowed to be rich; if they were, they would warp the rules set out by protocols of communication across settings of clinical relevance. Univocity rather than plurivocity; or, an easily curbed plurivocity.

In research and education, terms can be rendered subtle and argued over. Not so in the clinical context, where patients are to be 'managed'. There the terms must have the clarity required to meet their administrative

purposes and contexts of usage.[7] Clarke and Casper also make the point that 'a pattern of increasing complexification of classes of categories has been followed by simplification. This process has occurred as Pap smears have become more widely used as a screening tool nationally and internationally, increasingly institutionalized, and increasingly assessed epidemiologically and economically as an indicator for active prevention or early intervention. Simplification (such as a reduction in number of categories) is usually requisite for the kinds of standardization such multisite applications involve'[8] (Casper & Clarke 1998, 264).

TBS components

The Bethesda System has two main components: a statement of adequacy and a descriptive diagnosis (Solomon et al. 2002, 2116).[9]

(a) Sample adequacy

TBS better defined criteria for a good specimen from the squamocolumnar junction of the female cervix. An adequate specimen is one that is properly labelled, with relevant clinical information, properly fixed, and which shows microscopically an adequate number of well-preserved, evenly distributed evaluable cells reflecting an ectocervical and endocervical component. Following a review of the pertinent but admittedly sparse literature, the TBS criteria committee formulated the quantitative criteria for specimen adequacy. The Criteria Committee of the 1991 Bethesda System Workshop stated (National Cancer Institute Workshop 1993, 115–124)

> Well-preserved and well-visualized squamous epithelial cells should be spread over more than 10 percent of the slide surface. An adequate endocervical / transformation zone component should consist, at a minimum, of two clusters of well-preserved endocervical and / or squamous metaplastic cells, with each cluster composed of a minimum of at least five appropriate cells. This definition applies to specimens from both premenopausal and postmenopausal women with a cervix, except in the situation of marked atrophy in which metaplastic and endocervical

> cells often cannot be distinguished from parabasal cells. In cases of marked atrophic changes, the absence of an identifiable endocervical / transformation zone component does not affect the specimen adequacy categorization of a specimen otherwise determined to be satisfactory for evaluation.[10]

There should be at least 5000 cells per slide. This number is subject to alteration. Any specimen slide with abnormal cells will be described as satisfactory for evaluation regardless of the number of cells. (I would argue that this exemption from quantitative strictures is related to the issue of litigation attached to false negative readings.) In many instances, in the absence of data to provide a basis for establishing criteria, the committee drew on the judgment of its members. The criteria were an initial attempt to develop a more standardized approach to the evaluation of adequacy. The indicated percentages were to be used as general ranges, rather than as cut-offs.

Further, cells should not be too thickly smeared, not obscured by blood, mucus etc. To establish definitions of adequacy, the committee also addressed the issue of ambiguities in cellular collection, for instance, in the case of slides with cell clustering, atrophy or cytolysis which were difficult to count. The recommendation was that cases of borderline adequacy were to be addressed with professional judgement in hierarchical review. The increased stringency in the definitions of specimen adequacy also means that the number of unacceptable specimens will rise significantly. The unsatisfactory category includes specimen slides without sufficient cells for reliable interpretation as per TBS specifications.

The incorporation of specimen adequacy as an integral part of the report has been acknowledged as one of the most important contributions of TBS. The impact on laboratory practice has been dramatic. Surveys conducted by the College of American Pathologists revealed that in 1990, only 35 % of responding laboratories routinely reported specimen adequacy; by 1992, this figure had increased to 85 %.

The implications of the new and more stringent criteria of specimen adequacy can be analysed as follows. The new criteria sought to achieve non-artefactual reproducibility of samples and their readings across locales where they were to be collected, evaluated, interpreted, and acted upon.

The TBS criteria sought to extend to the practices of gynaecological cytology those criteria of adequacy that are sought after in the biomedical sciences based on microscopic anatomy. TBS and such 'reporting guidelines' limit artefactuality and idiosyncrasy by limiting the localisation of material and the language that makes that material as something definable, cognizable, recognizable, workable. It would be a mistake, I am convinced, to think that TBS is just a set of 'reporting guidelines'. How we say makes what we say; what we say makes what we do. Medium and message are not separable; although the scientific chant is that the truth remains the truth regardless of language. Truthsaying is truthmaking; language makes truthsaying possible; TBS provides new rules for new language games, new ways of truthsaying and truthmaking. TBS is an act of articulation that is also a means to further acts of articulation. TBS and what it makes say-able and do-able co-produce each other.

First, the specimen has to be from the right *topos* of the cervix. As the major purpose of the Pap smear is the earliest possible diagnosis of cervical cancer and its precursors, the most common site of origin for these lesions, the cervical squamocolumnar junction (transformation zone) is to be adequately sampled. Second, such a specimen ought to exhibit an adequate amount of well-preserved and properly-stained cellular material to be useful for diagnosis. It should remain preserved for indefinite slide storage and later review if warranted – to be an 'immutable mobile' in Bruno Latour's terms.

TBS connotes a reconfiguration and a consolidation of many different systems of heuristics and practice. TBS and such 'reporting guidelines' limit artefactuality and idiosyncrasy by limiting the localisation of material and the language that makes that material as something definable, cognizable, recognizable, workable. They 'manage' without seeking to end the tension between heuristics / experience / judgement and protocols. The Committee experts recognized that strict objective criteria cannot be used in an 'as is' way. But ambiguity, and terms of ambiguity were to be anticipated, pre-empted and sequestered as far as possible.

(b) Descriptive diagnoses

The nomenclature for the descriptive diagnoses in TBS incorporated the knowledge of the biology of cervical neoplasia and precursors developed over the past 40 years. Three broad categories were defined: benign cellular changes, squamous epithelial cell abnormalities and glandular cell abnormalities. The following discussion will focus on the categories of squamous epithelial cell abnormalities, short of carcinoma, particularly on ambiguous atypia.

Infections and reactive changes

Although the main purpose of cervicovaginal cytology screening has been the detection of cervical cancer precursors, reporting findings of infectious or reactive conditions expands the patient management potential of TBS. One net covers more surface – in a good example of the principle of parsimony. With attention to diagnostic criteria, excellent specificity can be achieved for the cytopathologic diagnosis of fungal elements *Trichomonas vaginalis*, and herpes simplex virus. In TBS 2001, cervical cytology specimens that contain no epithelial abnormalities are listed under the category 'negative for intraepithelial lesion or malignancy'. This category now encompasses the TBS 1991 categories of 'within normal limits' and 'benign cellular changes'. The presence of organisms such as *Trichomonas vaginalis* or those consistent with *Candida* will be included as a comment in this 'negative' category. Components that are optionally listed in the negative category include atrophy, radiation and inflammation.

The diagnosis of *Chlamydia* is not listed in TBS due to the acknowledged low diagnostic accuracy of routine cytology for this organism. We should take note here that human papillomavirus (HPV) is not included under the category of infection but instead is considered under the category of epithelial cell abnormalities. Cervical cancer has a well-documented infectious aetiology. The high-risk strains of HPV (especially types 16 and 18) are associated with the development of precancerous cervical lesions. Thus patients who are positive for HPV, specifically the high-risk strains, are triaged into the category of epithelial cell abnormalities for further examination and action.

Epithelial cells seen in cervical-vaginal specimens manifest reactive morphologic changes in response to a variety of traumatic insults, such as infection, inflammation, and radiation. It is important to recognize these benign reactive features to avoid over-interpretation and resulting false-positive diagnosis. Repair, radiation, atrophy, and intrauterine contraceptive devices can all induce cellular changes that may mimic intraepithelial lesions or even cancer. However, given an appropriate clinical setting, most of these samples have distinguishing morphologic features that enable a specific diagnosis.

Squamous cell abnormality

The category of squamous cell abnormality includes the subcategories of atypical squamous cells of undetermined significance (ASCUS), low-grade squamous intraepithelial lesion (LSIL), high-grade squamous intraepithelial lesion (HSIL), and squamous cell carcinoma (SCC).

Atypical Squamous Cells of Undetermined Significance (ASCUS)

The most common abnormal Pap test result is one of uncertainty, atypical squamous cells of undetermined significance (ASCUS). As Papanicolaou also observed (Papanicolaou 1954, 20–21)

> Smears cannot always be judged as positive or negative. There are cases in which cytologic findings are inconclusive. A classification taking into consideration the relatively large group of questionable smear findings is therefore necessary. One may often experience great difficulty in classifying cells that deviate from the normal type but show no malignant characteristics. An intermediate class between the entirely normal and the suspicious groups appears thus to be necessary. A similar need for subdivision exists in the positive group. There are instances in which the results are of an overwhelmingly positive character, leaving no doubt as to their final interpretation. On the other hand, there are cases in which there is strong but not fully convincing evidence of malignancy.[11]

In the United States, about 2 million ASCUS Pap reports are issued each year. Since TBS 1988, the category of 'Atypical Squamous Cells of Undetermined Significance' (ASCUS) has included cells for which a reliable

interpretation of SIL cannot be made although the cells contain features that are more marked than merely reactive changes. 'Undetermined Significance' means that a judgement cannot be made and that further triage is to be carried out.[12]

TBS 1991 did not qualify ASCUS, but it was suggested that it be qualified e.g., ASCUS favour reactive, or ASCUS favour neoplastic. But there was no consensus on this sort of definition and there was no reproduction of these qualifiers across locales. A comprehensive study showed that the ASCUS cytology report is not reproducible, even among expert cervical cytopathologists.

Other research demonstrates that while the majority of ASCUS Pap test results reflect a benign reactive process, 5 % to 10 % of the women with ASCUS results harbour underlying high-grade squamous intraepithelial lesions (HSILs). Subsequently, evidence was adduced to show that ASCUS was not to be ignored. One study on 4143 diagnoses of ASCUS with subsequent histology reported that in 63 per cent of the women, SIL or malignancy was detected on further follow up (Jones & Novis, 1996).[13] Another study of more than 46,000 women receiving routine cervical cytology screening demonstrated that the most common cytology result immediately preceding the diagnosis of histologic HSIL or greater was ASCUS. (39 per cent of all cases of HSIL or cancer). Let us note here that such judgements could be made because ASCUS was already a defined category of non-definition.[14] Studies to establish the need to define and further triage ASCUS exemplify a co-productive 'convergence' between knowledge and infrastructure, the ways of 'disciplining of representational practice' in accordance with a standardized classification system (Star et al 1997).[15]

These findings were related to a renewed focus on ASCUS in TBS 2001. The new category of 'Atypical Squamous Cells' emphasizes the importance of quelling ambiguity and determining risk status by dividing smears into likelihood of being 'negative' or harbouring an SIL. ASC was divided into two sub-categories. 'Atypical Squamous Cells of Undetermined Significance' (ASC-US) and 'Atypical Squamous Cells – Cannot Exclude HSIL', or ASC-H. TBS 2001 eliminates 'ASCUS favour reactive' with the notion that most of the specimens in this category will be down-

graded to 'negative'. ASC-US includes most cytology results previously also categorized as ASCUS-NOS (Not otherwise specified) or ASCUS, favour SIL. ASC-US excludes cytology suggestive of HSIL. ASC-H, around 5–10 per cent of all ASC, is interpreted as cytological changes that suggest HSIL but cannot definitively be interpreted as HSIL. The positive predictive value of CIN 2, 3 in ASC-H is higher than in ASC-US but not as high as in HSIL.

LSIL / HSIL

Next up after ASC comes LSIL (same as in TBS 1991). It includes the categories of human papillomavirus, mild dysplasia and CIN 1. There was a debate in TBS 2001 as to whether CIN 2 was to be put in LSIL or HSIL. This was resolved with the decision that CIN 1 stayed in LSIL and CIN 2 in HSIL. The rate of LSIL is more variable than the rate of HSIL. The accuracy rate of interpretation of LSIL is 80 per cent. 20 per cent are estimated to be misclassified HSIL.

The category of HSIL remains unchanged from 1988 to TBS 2001 and includes moderate dysplasia, severe dysplasia, and carcinoma in situ, which were once independent categories, another instance of simplification by clubbing together.

ASC-US/H, triage, and the pre-emption of legal risk

The definition and clinical significance of ASCUS is one of the most controversial areas in gynaecological cytology. This category is an attempt by the classifiers to come to grips with a messiness, which Star and Bowker call 'residuality', something that resists the desired neatness of a standardized classification system and thereby creates a strain for the classification system (Star & Bowker 2007, 274).[16] ASCUS is not a single diagnostic entity, but rather encompasses a spectrum of cellular changes and reflects a variety of pathological processes that cannot be more specifically categorized. The diagnosis of ASCUS reflects the discretionary and judgement-based contingent reality of the work of classifying cytological changes. The cellular changes seen in a cervical smear constitute a spectrum of change.

A diagnosis is based on the cellular criteria of multiple cells as interpreted by the human eye. In many cases, the diagnosis may be in a grey zone between categories as the cells show morphological features that suggest more than one category.[17]

The work of slotting things right is crucial in the management of risk and uncertainty. 'Nature' is to be carved at the joints; this means two types of work that rely on and co-produce each other. On the one hand, the work of identifying, howsoever tentatively, risk boundaries / margins, and on the other, designing and initiating procedures to populate these zones with risky content. This is the work of constituting a zone of risk *and* deciding the limina / checkpoints at which the risk starts, stops, progresses etc. To decide on the contents of risk zones is also to make and to refurbish the boundaries, the thresholds / decision making points. Diagnosis is very much a *rite de passage*.

The absence of pathological evidence is not to be read as evidence of pathological absence. This is where a category like ASC is useful to the community of medical practitioners and we will see how it allows triage interventions that are pre-emptive of both oncogenic and legal risk. In effect, ASC judgements, whether ASC-US or ASC-H, mean that the patients so classed must be monitored closely, 'managed' and 'surveilled'. The difference in targeting women differently categorized is a matter of degree, not kind. In fact, it may be argued, and some authors do argue, that ASC-US judgements offer even more room than ASC-H for such surveillance, to excess, because of the inherent ambiguity of the categorization. Since TBS was published and implemented, many clinicians have felt overwhelmed with ASCUS diagnoses in their patient populations. The reasons underlying this ASCUS explosion may be linked to a constant threat of litigation against false-negatives.[18]

The atypia that are slotted in ASC-US and in ASC-H are not themselves clearly defined by the norms that define benignity of lesions or malignancy. As abnormalities, their identifiability inheres in the norms that identify the contents of the categories that lie to either side of them. But those norms are not *per se* clearly applicable to ASC themselves. The contents of ASC are identifiable only by default – by the identifiability of the benign or the malign, which make a frame for this grey zone.

ASC-US/H is an ambiguity that seeks to avoid the risk of litigation that attaches to false-negative diagnoses; placing the patient so assessed in a zone of irresoluble risk to avoid another sort of risk. Here is a case of a medically defined and maintained ambiguity that ostensibly protects the medical practitioner from the legal risk of making a false definitive statement. Since the practitioner cannot also err on the side of caution, by making a false-positive diagnosis and putting the patient through a battery of (overtly) aggressive management, here comes a way to define, class, and monitor the (not-quite-yet) patient.

ASCUS and such categories are like DMZs or hotel lobbies; they can be as populous or as deserted as possible. They are necessary ambiguities. Something like ASCUS is preserved because it makes it possible to keep things in suspension till the rites of decision are completed, till the patient can be sent home for a bit till the next check up or right on into the cancer ward. ASCUS categories are also the point where the cytolab and the clinician can negotiate / suspend their transactions over the patient's body. I might also remark here that the creation of such 'holding pens' is a fairly standard way for the institutions of modernity to deal with ambiguities. Zygmunt Bauman, following Simmel, has written at length on the 'anthropoemic' or the 'anthropophagic' ways in which modern European society has dealt with people who cannot be classed, who are defined as existing in a liminal zone, as neither here nor there. One might similarly be stuck forever in an international airline terminal or one could be a refugee in a camp awaiting a decision on asylum or deportation. ASC-US judgements place the patient in a zone of absolute liminality; they create a situation where it becomes the patient's responsibility to have follow-up checks done, because the clinician has already been circumspect enough to serve a warning. Responsibilization moves forward because the patient would obviously want to be 'down-graded.' Follow-up monitoring is the modality to this end. ASC judgements expand the definition of at-risk and risky populations of women, to be surveilled and managed. The expansion of definition is accompanied by a quasi-consent-based and institutionalised insistence that women 'comply' with the system of follow-up checks. Shame, warning, coercion – covertly applied through reminder systems. Those who do not report for monitoring check-ups

are called 'non-compliant'. There is stigma in non-compliance; possible stigma in compliance; certain stigma in being diagnosed with something.

ASCUS judgements are directly connected to a legal system; it is interesting that the legal system to which false-negative patients take recourse is also the avenue that leads to categories such as ASCUS/H that lead to more monitoring and surveillance of the female populations, not treating them or curing them or even outright pathologizing them, but placing them in a semantic-diagnostic-conceptual-iatrological 'holding pen'. And in the name of health and responsibility of the population, this step is also placed beyond critique. People want to be well, but they want even more to be judged as well, healthy, safe. The dimensions of disease are manifestational, etiologic, and mechanistic. To adapt from Talcott Parsons, well-healthy-safe :: illness-disease-sickness :: personal-medical-legal. Since cervical cancer is often asymptomatic till advanced stages, the manifestational aspect is shifted deeper down to the microscopic level with the smear. The HPV etiology is then appended to the ASCUS manifestation, using the TBS lexicon, to make ASCUS less polythetic, less dubitable as evidence of underlying pathology.

A primary goal of Pap screening and follow-up procedures is to prevent cervical cancer by identifying and treating high-grade precursor lesions. Given the relatively large proportion of HSIL cases that are associated with ASCUS cytology, effective triage of ASCUS reports is essential. Although routine colposcopic evaluation of all ASCUS cases would provide the greatest patient protection, the frequency of ASCUS makes this impractical. Recent guidelines propose immediate colposcopy, repeat Pap testing, or adjunctive testing for triage of ASCUS cytology. Researchers have sought diagnostic tools to identify higher-risk ASCUS cases and direct them to colposcopy and in turn, refer those at lower risk for periodic repeat Pap testing.

Several studies suggest that the Hybrid Capture test (Digene Corporation, Beltsville, Md) for cancer-associated HPV types is useful for identifying patients with ASCUS Pap results who have underlying HSIL. The requirement of a second examination to repeat Pap testing or obtain a specimen for HPV testing raised concerns about the efficiency and cost

effectiveness of these approaches for ASCUS triage. Recent pilot studies using liquid-based cytology (ThinPrep Pap, Cytyc Corporation, Boxborough, Mass) demonstrated the utility of a single specimen for cytological diagnosis and subsequent HPV testing, when indicated. If HPV DNA testing were conducted on residual material collected for routine liquid-based Pap screening, then the second examination could be eliminated and ASCUS patient management decisions could be streamlined.

Triage: why is this word of mixed agricultural and military provenance used in the context of HPV testing to sort out ASC readings into low and high risk groupings?

According to Diane Solomon, in the context of cervical cancer screening, (Solomon 2003, 97–98)

> Triage is an additional step interposed between screening and diagnosis to further stratify individuals with positive primary screening results according to risk for the disease state. (…) The utility of a triage test in the context of a screening program will depend not only on the performance characteristics of the test itself but also on the target screening population, the prevalence of disease, the screening test employed, the costs of follow-up, the available resources (logistical and monetary), and patient compliance. Triage is of most value when the screening test lacks specificity and/or the diagnostic procedure is expensive or a limited resource. An efficient triage test should reduce overtreatment, patient anxiety and inconvenience, and overall management costs, usually by reducing the number of diagnostic procedures performed – all without sacrificing sensitivity for detection of disease. However, if the initial screening test is highly specific and/or the triage test is positive in the vast majority of cases and/or the triage test results in a decrement of sensitivity, then the triage test has questionable utility. Cost-effectiveness analyses are critical to evaluating the interplay of population prevalence, sensitivity for disease detection, specificity, and overall management costs with various screening and triage strategies.[19]

Iserson and Moskop also explicate the relationship of triage to questions of the scarcity of resources (2007, 275–276)

> Though 'triage' may be used in an extended sense to refer to any decision about allocation of a scarce medical resource, we believe that use of the term in its primary sense requires that 3 conditions be satisfied:

1. At least a modest scarcity of health care resources exists. (…)
2. A health care worker (often called a 'triage officer') assesses each patient's medical needs, usually based on a brief examination. This assessment distinguishes the practice of triage, in which microallocation decisions are made about specific individuals according to face-to-face encounters, from the process of macroallocation, such as decisions made by legislators or administrators when allocating health care funds or other resources to different population groups.
3. The triage officer uses an established system or plan, usually based on an algorithm or a set of criteria, to determine a specific treatment or treatment priority for each patient. This condition distinguishes triage from purely ad hoc or arbitrary decisions about distribution of health care resources. (…)[20]

Triage is a process, and a complex of decision making acts and events, implicated in the making of a resource-beneficiary ratio. The sorting out of populations into categories that are the targets of different interventions is related to the key question of resource management. Resources refer, on the one hand, to the time, space, money, personnel, equipment involved in clinical settings and directed at the populations interned there. On the other hand, the patient is also seen as a resource whose withdrawal from 'normal' life-settings diminishes his/her value as a resource. Such withdrawals are therefore closely monitored. The high-risk groups are sorted out and withdrawn out of normal life-worlds and settings into treatment scenarios, and the scarce resources of hospital care are applied to them.

Patient flows are managed just as are the health resources applied to them; the critical point to be made here is that each so-called beneficiary or do-gooder is also a resource, an actor, and a mediator in the network, in the flow, in the configuration. With the growing insistence that cervical cancer has a dynamic rather than progressive staging, the valuation of time and triage increases proportionally. Note here that the TBS reporting format must take into account not only spatial (cervical transformation zone, localised lesions versus metastatic spread) but also temporal variables (cancer staging) (Star & Bowker 1996).[21] Time, money and the law are found in a triadic relation, exemplified in the case of cervical cancer screening and triage.

The question of triage and streamlining is also connected to the inverse relationship of visual fatigue and interpretive accuracy in smear analysis, and has a well-defined medico-legal angle that has had important

ramifications for the profession of the cytotechnologist. The Pap smear is seen as an imperfect screening procedure and has a significant error rate. The scientific literature is replete with studies that report these error rates expressed in terms of false-negative rate (sensitivity) and false-positive rate (specificity).[22] The average Pap smear contains 300,000 events, including epithelial cells, inflammatory cells, and microorganisms. After examining every microscopic field, a report must be rendered that includes a statement of adequacy, documentation of cell types, documentation of epithelial cell abnormalities, and presence of microorganisms, including abnormal bacteria, fungi, and protozoa. Slide interpretation must also consider available medical history, including results of previous Pap smears.

This connects to the question of the human resources for screening. The demand for experienced cytotechnologists exceeds the numbers available and the shortfall is not likely to decrease given the ever-increasing stress on cancer screening. Cytology screening is demanding not only in terms of performance, but also in the application of the regulations and standards to this litigation-prone professional field. In the US, regulations applicable to clinical laboratories became effective on 1 September 1992, implementing provisions of the Clinical Laboratory Improvement Act of 1988, Public Law 100–578. Such regulations were a response to the Pap mills that churned out unregulated reports, replete with false negatives, with cytologists, often untrained, working from home, and the fact that cancer screening had become big business (Singleton et al. 1995).[23]

The number of slides handled in a laboratory, the demands for wider screening, the life-stage monitoring of the women from whom the samples come and whose lives are ordered to a variable extent around the screening to which they comply and by which they are 'managed', the question of experience and visual efficiency or fatigue of the technologists, the regulations of the profession of cytotechnology and also the federal regulations in the US dictating best practice conditions for smear analysis, the clinicians who work on the basis of the smear reports to define further courses of action for their patients – all are linked in a complex array of social worlds, arenas marked by interpretive struggles and varying interests. A configuration – a fish-scale overlapping of economies – takes shape and is consolidated in fits and starts.

We see the coimbrications of: an economy of credibility (adequate smears and non-artefactual smear readings); an economy of tools of credible knowledge; an economy of the targets-tools-beneficiaries of the economy of credible smears (intensive surveillance rather than repeated visits – 2 tests in 1 sample with liquid cytology); parsimony and economy of the taxonomy of pathology; parsimonious use of diagnostic and therapeutic time-frames, of screening-diagnostic (colposcopy / biopsy) resources, of therapeutic interventions (loop electrical excision procedure, conization); parsimony and economy of language (less risk of litigation over disagreements, less money and time spent on back-and-forth bickering). As cytologists-clinicians-scientists-patients come to share a lexicon of terms with more rather than less mutual intelligibility, their social worlds are tied together in a complex convergence of economies marked by concern for the parsimonious use of resources.

Conclusion

The complex of acts that animate, institutionalize and maybe blackbox guidelines such as TBS are an obligatory passage point for the patients, cytologists and clinicians, as also the state agencies and the families of those affected and all the actors involved in the CCS bandwagon. TBS was possible and perhaps adopted the way it was because the organization of reportage of Pap smears had not yet been uniformly conventionalized. So the bandwagon that was initiated in 1988 and was consolidated thereafter was able to develop and incrementally revise and authorize the TBS package. After 1988, as the bandwagon grew and the TBS package spread across cyto-clinician networks, 1991 and 2001 saw more people climbing aboard that particular bandwagon and adding to the package (Fujimura 1988).[24]

TBS enables a communicative interface between social worlds. The use of 'reporting' guidelines such as TBS forms cell samples into scripts that can move with a suitably curtailed plurivocity from doctors to cytolabs and back, and then also from the teaching hospital to the research laboratory. TBS is part of the making of shared epistemic spaces, collaborations

between cytologists and clinicians, and the regulation of information flows between different professional camps. These flows are marked by the simultaneously collusive and adversarial elements of the relationships between these camps.

Drawing from Fujimura's analysis of the cancer research bandwagon, TBS is a highly transportable theory-method package and it makes its object – Pap smears – similarly transportable (Fujimura 1988).[25] Something transportable and relatively unequivocal as the TBS package makes sense in biology as big as the cancer research-therapy-diagnostics complex.(Bud 1978).[26] Unequivocally diagnosed samples means non-litigious treatments; they also mean non-artefactual research samples; they also mean suitable targets for experimental drugs.

TBS is protean and robust like any respectable boundary object – package; passage point; conduit; interface. TBS is mobile (though not an immutable mobile), it transports stuff, and as a conduit it is itself a state-of-suspension. (The Pap smear is, from the start, a nascent script that is put through interpretative procedures in the TBS.) If we say that the TBS package of diagnostic terms enables the transportation of Pap smears and reports on Pap, it serves as an interface between worlds of practitioners – cytologists, clinicians, and also the researchers who use tissue samples for their work – which means that communication takes place at this interface. We may view TBS as playing a role akin to that of the International Classification of Disease. In Star, Bowker and Neumann's discussion of the ICD, it emerges that this package is far more than a classificatory scheme. The ICD, like the TBS marks several levels of 'inter-adjustment'; between social worlds (such as epidemiologists, health officials, health insurance companies, census officials), and between social worlds and information artifacts (Star et al. 1997).[27]

But as an interface, TBS is also a transit zone. This means that TBS is marked by the intractable ambiguity of its necessary burden of residual categories. So also is the Pap smear which is made and read through the TBS lexicon. The guidelines seek to define and curb ambivalence and ambiguity, but that means that ambivalence and ambiguity must also be sustained. Some sentences have to be left with ellipses and question marks at the end. It is through that sustenance of the ambiguity of a

boundary zone that the stability of the boundary is maintained; those whose roles are sustained in the networks of cancer screening also draw on that ambivalence for the longevity of their roles and actions in the network. Instabilities are permanent parts of the network (Clarke & Casper 1998, 277).[28] Both the guidelines and the actors who contend with these guidelines (making and remaking and sustaining them in their use and implementation) are marked and maintained by the tension between the practical reasoning in experience and judgement and the heuristics-in-protocols.

Acknowledgements

I am grateful to Dr Günter Getzinger and Dr Christine Wächter for their involvement and interest in this work. My colleagues at the IAS-STS, in particular Dr Andrea zur Nieden and Dr Aaro Tupasela have been very forthcoming with their scholarly advice and inputs towards the ideas I have tried to outline in this paper. I remain indebted to them.

The paper has been completed as part of my ÖAD-funded research project 'Biomarkers for cervical malignancies: Discourses of risk and gender in gynaecological diagnostics', from March 2008 to February 2009, at the IAS-STS, Graz, Austria.

Notes and References

1 Papanicolaou, G. N. (1954), *Atlas of Exfoliative Cytology*, Cambridge, MA: Harvard University Press.

2 Clarke, Adele E. and Casper, Monica J (1996), 'From simple technology to complex arena: Classification of Pap smears, 1917–1990', *Medical Anthropology Quarterly* 10 (4): 601–623.

3 National Cancer Institute Workshop a (1989), 'The 1988 Bethesda System for reporting cervical / vaginal cytological diagnoses', *JAMA* 262: 931–4; National Cancer Institute Workshop b (1993), 'The 1991 Bethesda System for reporting cervical / vaginal cytologic diagnoses', *Diagn Cytopatho* 9: 235–246.

4 National Cancer Institute (a) 'Workshop Recommendations', NCI Bethesda System 2001 http://bethesda2001.cancer.gov/postwrkshp_recs.html [15 August 2008]; National Cancer Institute (b) 'Workshop Forum Groups (Committees)', NCI Bethesda System 2001 http://bethesda2001.cancer.gov/committees.html

[15 August 2008]; National Cancer Institute (c) 'Bethesda System 2001 Terminology', NCI Bethesda System 2001 Workshop http://bethesda2001.cancer.gov/terminology.pdf [15 August 2008].

5 Cox, J. T. (1997), 'ASCCP practice guidelines: Management issues related to quality of the smear', *J Lower Genital Tract Disease* 2: 100–106.

6 Casper, Monica J. and Clarke, Adele E. (1998), 'Making the Pap smear into the right tool for the job: Cervical cancer screening in the USA, circa 1940–95', *Social Studies of Science* 28 (2): 255–290.

7 Singleton, Vicky, and Michael, Mike (1993), 'Actor networks and ambivalence: General practitioners in the UK cervical screening programme', *Social Studies of Science* 23 (2): 227–264. See this paper for a fascinating point: How the system insists that practitioners remain dependent on cytological analysis, remain ignorant of reading the Paps for themselves, cannot be critical of the networks of which they are a part, and thereby help to blackbox the Pap smear collection and analysis system.

8 Casper, Monica, J. and Clarke, Adele E. (1998), see note 6.

9 Solomon, D., Davey, D., Kurman, R., Moriarty, A., O'Connor, D., Prey, M., Raab, S., Sherman, M., Wilbur, D., Wright, Jr, T., Young, N., for the Forum Group Members and the Bethesda 2001 Workshop (2002), 'The 2001 Bethesda System: Terminology for reporting results of cervical cytology,' *JAMA* 287: 2114–2119.

10 National Cancer Institute Workshop (1993), 'The Bethesda System for reporting cervical / vaginal cytologic diagnoses', *Acta Cytol* 37: 115–124.

11 Papanicolaou, G. N. (1954), see note 1.

12 Kurman, R. J., Henson, D. E., Herbst, A. L., Noller, K. L., and Schiffman, M. H. (1994), 'Interim Guidelines for Management of Abnormal Cervical Cytology', *JAMA* 271 (23): 1866–1869. The authors outline several options for clinical follow-up of ASCUS. These include; (1) follow-up by repeat smears every 4–6 months for 2 years until there have been 3 consecutive negative smears or (2) colposcopy if the patient is considered at high risk, if another abnormal smear is reported in the follow-up period, or if poor compliance is anticipated. If ASCUS is associated with severe inflammation and a specific infectious agent is identified, re-evaluation should be performed after appropriate treatment. If ASCUS is associated with atrophy, a repeat smear should follow a short course of estrogen. Ancillary diagnostic techniques, such as HPV typing, may prove to have utility in the identification of women who harbour occult HSIL at the time of the ASCUS diagnosis or possibly predict which low-grade lesions will progress over time to HSIL.

13 Jones B. A. and Novis, D. A. (1996), 'Cervical biopsy-cytology correlation: A College of American Pathologists Q-Probes study of 22439 correlations in 348 laboratories', *Arch Pathol Lab Med* 120: 523–531.

14 Schiffman, Mark and Solomon, Diane (2003), 'Findings to date from the ASCUS-LSIL Triage Study (ALTS)', *Arch Pathol Lab Med* 127: 946–949. 'Controversy exists in the United States regarding the proper evaluation and management of low-grade squamous intraepithelial lesion (LSIL) and equivocal (atypical squamous cells of undetermined significance [ASCUS, now ASC-US]) cervical cytologic interpretations. To address this issue, the National Cancer Institute initiated the ASCUS LSIL Triage Study (ALTS). ALTS is a multicenter, randomized clinical trial designed to evaluate 3 alternative methods of management, namely, immediate colposcopy, cytologic follow-up, and triage by human papillomavirus (HPV) DNA testing. This article summarizes the major findings of ALTS that have been published to date. Patients with ASCUS (n= 3488) or LSIL (n = 1572) were randomly assigned to research arms between November 1996 and December 1998, and were monitored for 2 years. The disease outcome was histologic cervical intraepithelial neoplasia (CIN) 3/cancer. The prevalence of oncogenic HPV was too high to permit effective triage of LSIL using HPV DNA testing by Hybrid Capture 2. However, for the women referred with a cytologic interpretation of ASCUS, HPV triage proved useful, with sensitivity equivalent to immediate colposcopy and a halving of colposcopic referrals. Among older women with ASCUS, HPV testing remained sensitive for detecting CIN 3 and cancer, but the referral percentage was dramatically lower compared to younger women. ALTS yielded insight into the performance of cytology and histopathology; experienced pathologists differed significantly in their interpretations of cervical abnormalities, especially histologic CIN 1 and cytologic ASCUS. Nonetheless, it was possible to distinguish a relatively uncommon type of ASCUS, equivocal for high-grade squamous intraepithelial lesion, that has a high positive predictive value for identifying women with underlying high-grade CIN. Many additional analyses are underway.' (Abstract)

15 Star, Susan L., Bowker, Geoffrey C., and Neumann, Laura J. (1997), 'Transparency at different levels of scale: Convergence between information artifacts and social worlds,' http://epl.scu.edu:16080/~gbowker/converge.html [1 November 2008].

16 Star, Susan L. and Bowker, Geoffrey C. (2007), 'Enacting silence: Residual categories as a challenge for ethics, information systems, and communication', *Ethics and Information Technology* 9: 273–280. The authors discuss the ways in which a category exhibits the aspects of 'residuality'. An object may be residual if it 'embodies two or more of the categories in the category schema, or falls between the cracks, or otherwise misfits (with varying range of misfit as a variable) the pigeon-holes (…)' (274).

17 See Doeberitz, Magnus von Knebel (2006), 'Biomarkers in screening for cervical cancer', in Monsonégo, Joseph (Ed.), *Emerging Issues on HPV Infections: From Science to Practice*, Switzerland: Karger, 1–19. The author says 'On the technical level, the subjective interpretations of scoring criteria that are used to classify the cytological samples were and still are associated with a significant lack of reproducibility of test results. This causes substantial secondary costs due to repeated testing and further clinical work-up of patients with ambiguous test results and may also lead to psychological as well as physical distress of the affected women. Due to these limitations of an otherwise very successful concept in cancer prevention, substantial research efforts were undertaken to better understand the molecular events involved in cervical carcinogenesis and to delineate novel screening methods (…)' (1). In other words, the current ambiguities of visual histologic analysis necessitate technical refinements, such as by the use of diagnostic tools such as biomarkers for disease progression. The definition of residual categories such as ASCUS thus proves to be epistemically profitable.

18 See Cox, J. T (2006), 'HPV testing in patient management: Atypical squamous cells of undetermined significance and low grade squamous intraepithelial lesion', in Monsonégo, Joseph (Ed.), *Emerging Issues on HPV Infections: From Science to Practice*, Switzerland: Karger, 120–139. ASCUS figures are lower in countries with lower patient vs. doctor litigation statistics. At present, 2.5 million out of 55 million (approximately) Pap smears processed per annum in the US are rated ASC-US, as contrasted to 1.2 million rated as ASC-H.

19 Solomon, Diane (2003), 'Role of triage testing in cervical cancer screening', *JNCI Monographs* 31: 97–101.

20 Iserson, Kenneth V. and Moskop, John C. (2007), 'Triage in medicine, part I: Concept, history, and types', *Annals of Emergency Medicine* 49 (3): 275–281.

21 Star, Susan L. and Bowker, Geoffrey C. (1996), 'Of lungs and lungers: The classified story of tuberculosis', http://epl.scu.edu:16080/~gbowker/tb.html [1 November, 2008]. The authors say 'Thus, any classification system should include both spatial and temporal dimensions, but standardized classifications tend to emphasize space alone. That is, classifications are rarely developmental, and often presented as spatial demographic distributions. Even where stages of a disease may be categorized, these stages are abstracted away from biographical continuity and more subtle temporal issues. As the problems of time emerge in the lives of patients and the work of classifiers, those spatial compartments break down in interesting ways: a formal hierarchy of mutually exclusive categories becomes a set of overlapping contradictory classes.' This points to the fact that the neatness of the formal taxonomy in TBS is confounded by the dynamism of various in vivo factors; thus the residual ambiguities of smear atypia.

22 Depending on the study cited, the reported false-negative rate ranges from six to 55 percent. See Gay J. D., Donaldson, L. D., and Goellner, J. R. (1985), 'False negative results in cervical cytologic studies', *Acta Cytol* 29: 1043–1046. The authors reviewed Pap smears from 339 patients with either histologically confirmed carcinoma in situ or invasive cancer and found 66 cases (20 percent) that had been reported as negative. After reviewing the original slides, it was determined that sampling errors by the clinician accounted for 62 percent of the false-negatives, interpretive errors by the pathologist accounted for 22 percent, and screening errors by the cytotechnologist accounted for 16 percent.

23 Singleton, Hugh M., Patrick, Roman L., Johnston, William W., and Smith, Robert A. (1995), 'The current status of the Papanicolaou Smear', *CA Cancer J Clin* 45: 305–320; Jones, Bruce A. and Davey, Diane D. (2000), 'Quality management in gynecologic cytology using interlaboratory comparison', *Archives of Pathology and Laboratory Medicine* 124 (5): 672–681. Some of the major provisions of the CLIA relating to cytology include the following: interpretations should not be reported on unsatisfactory specimens; a cytotechnologist is limited to a maximum of 100 slides per 24 hours (not to be used as a performance target) irrespective of the site of employment or the laboratory; the maximum allowable number of slides must not be completed in less than eight hours; if the individual works less than eight hours, the number of slides screened must be prorated; individual workload limits must be set for each cytotechnologist at least every six months, based on performance criteria established by the laboratory; all smears interpreted to be showing reactive or reparative changes, atypical squamous or glandular cells of undetermined significance, dysplasia, or malignancy must be examined by a pathologist; all non-gynaecological preparations must be examined by a pathologist; the slide examination performance of each cytotechnologist must be evaluated and documented through the re-examination of normal cases and feedback on abnormal cases; at least 10 percent of the normal cases must be re-screened; and the re-screened cases must include randomly selected normal cases and those from patients identified at high risk for developing cervical cancer. The CLIA also requires compilation of statistical data including the number of cases where cytology and available subsequent biopsies differ, where histology results were unavailable for comparison, cases in each descriptive category, and data on all screening discrepancies.

24 Fujimura, Joan H. (1988), 'The molecular biological bandwagon in cancer research: Where social worlds meet', *Social Problems* 35 (3): 261–283. 'A scientific bandwagon exists when large numbers of people, laboratories and organizations commit their resources to one approach to a problem'; 'A package of theory and technology is a clearly defined set of conventions for action that helps reduce reliance on discretion and trial-and-error procedures' (261). 'Changing conven-

tionalized and embedded work organizations involves a lot of convincing and persuading, buying and adopting, teaching and learning. Conceptual change in science in turn is based in individual and collective changes in the way scientists organize their work' (261).

25 Fujimura, Joan H. (1988), 'When tasks and procedures are standardized, that is conventionalized and routinized (…) they reduce the amount of tacit knowledge, discretionary decision making and trial-and-error procedures needed to solve problems. What is done to which material for what reason or purpose and with what outcome are all built into the "black box" of transportable technologies' (278). See note 24.

26 Bud, R. F. (1978), 'Strategy in American cancer research after World War II: A case study', *Social Studies of Science* 8 (4): 425–459. The author speaks of the rise of cancer research institutions post WW2 in the US. He discusses the rise in importance of the NCI as a funding agency, the declaration of the 'war on cancer' and the links between corporate interests of pharmaceuticals and the research-cum-experimental treatment hospitals such as the Sloan-Kettering.

27 Star, Susan L., Bowker, Geoffrey C. and Neumann, Laura J. (1997), 'Transparency at different levels of scale: Convergence between information artifacts and social worlds', http://epl.scu.edu:16080/~gbowker/converge.html [1 November 2008].

28 Casper, Monica, J. and Clarke, Adele E. (1998), 'Making the Pap smear into the right tool for the job: Cervical cancer screening in the USA, circa 1940-95', *Social Studies of Science* 28 (2): 255–290.