The Development and Evaluation of two Computer-based diagnostic aids in the field of Inherited Skeletal dysplasias and Malformation syndromes

Christine Margaret Hall

Great Ormond Street Children's Hospital and Institute of Child Health, London University

Submitted for Doctorate of Medicine (MD)

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

A retrospective review of all skeletal dysplasias referred in a period of one year for a diagnostic opinion, identified that, of all referring clinicians, general radiologists in particular had a low diagnostic accuracy.

Two computerised systems were developed. The first was an expert, knowledge-based system in which the knowledge of an experienced paediatric radiologist in the field of skeletal dysplasias was captured and diagnostic reasoning pathways defined. The second comprised a database of radiographic images with their radiological findings, of patients with skeletal dysplasias, combined with a powerful search facility for matching findings and images.

Each system has been tested individually in a standardised clinical trial by comparison with standard diagnostic methods used by radiologists, that is, by referring to lists of gamuts, standard reference textbooks and journals. Because the trial protocols and material were identical, a comparison of the relative value of each system could be made.

The results of the two trials showed that the use of each computerised system individually achieved significantly improved levels of diagnostic accuracy when compared with standard methods of diagnosis. Of the two systems the expert system showed slightly improved diagnostic accuracy compared to the image database.

The development and use of such systems will help to improve diagnostic accuracy and the quality of genetic advice given to patients and their families and patient management and treatment in difficult medical domains.

Table of Contents

Conjoint work statement

My thesis concerns the development and evaluation of two computer-based systems.

The first system involved artificial intelligence. My input has been as a radiologist identifying radiological abnormalities and defining and exploring diagnostic reasoning pathways. I am not a computer scientist. The computer scientists involved in creating the program were Professors John Washbrook and Elpida Keravnou of UCL Department of Computer Science. From my knowledge of skeletal dysplasias I was responsible for the medical input, in particular creating the 'dysplasia frames' entering data about each condition.

I was responsible for testing the diagnostic expert system in a clinical setting.

The second system was a database of radiographs and their findings. The data was all selected from my database of patients with skeletal dysplasias and the images digitised by me. I reported each image in a standardised format, identifying several findings associated with each image. I was not responsible for developing the software for the extensive search functionality although I identified the functions which could be helpful to medical users. The software was developed both by Oxford University Press and John Washbrook. It led to the publication of a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias (REAMS) in 2000, which I am submitting as Appendix D of my thesis.

I was responsible for testing the database in a clinical setting.

Acknowledgements

I would like to thank my colleagues in the Department of Computer Science at University College, London, Professors John Washbrook and Elpida Keravnou (now of the University of Cyprus) who were responsible for developing the artificial intelligence aspects of the expert knowledge-based system and for some of the search facilities used by the image database. I am grateful to Professor Andrew Todd-Pokropek for advice on image capture and quality, for the image database, and also to Professor David for advice on the trial protocol for the two systems, both of University College, London.

I would also like to thank my paediatric radiology colleagues – all experts in the field of skeletal dysplasias, for the time devoted to verification of the trial data. They are Dr Donald Shaw of Great Ormond Street Hospital, London, Professor Andres Giedion of Zurich, Switzerland and Professor Alan Oestreich of Cincinnati, Ohio. In addition Dr Shaw provided confirmatory evidence relating to diagnostic reasoning pathways when arriving at a diagnosis, for development of the expert system.

I am grateful to the Leverhulme Trust for the grant provided for development of the expert system, and to Oxford University Press for providing financial support for creating the image database and its publication as REAMS – a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias. Without their assistance this work would not have been possible.

My experience in this difficult and limited radiological field would not have been possible without the collaboration of my radiology and genetic colleagues at Great Ormond Street Children's Hospital. None of this would have been possible without the experience of the many patients suffering from these rare disorders and to them I am especially grateful. I hope my work in some measure has helped to improve the quality of their lives.

Finally I would like to thank Professor Robin Winter for years of collaboration and, as my principal supervisor, for providing critical evaluation and encouragement. He is sadly missed

Abbreviations

ADA	adenosine deaminase
AP	antero-posterior
C	cervical
DMC	Dyggve-Melchior-Clausen disease
ESDN	European Skeletal Dysplasia Network
FRCR	Fellow of the Royal College of Radiologists
LDDB	London dysmorphology database
MCD	metaphyseal chondrodysplasia
MED	multiple epiphyseal dysplasia
MPS	mucopolysaccharidoses
OPD	oto-palatal-digital
OSSUM	a radiographic database linked to POSSUM
POSSUM	Pictures of Standard Syndromes and Undiagnosed Malformations
REAMS	a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias
SED	spondyloepiphyseal dysplasia
SEDC	spondyloepiphyseal dysplasia congenita
SEDT	spondyloepiphyseal dysplasia tarda
UCL	University College London
UK	United Kingdom
USA	United States of America
ZW	Zweymuller-Weissenbacher

Chapter 1. Introduction

Skeletal dysplasias consist of a group of disorders in which there is a generalised abnormality of bone and cartilage growth and development. They are genetically determined, may present at any stage and continue to evolve as a result of active gene involvement throughout the life of an affected individual. The latest International Classification and Nosology of Constitutional Disorders of Bone (2001) [1] recognises about 250 skeletal dysplasias. This does not include reports of isolated cases.

Malformation syndromes are conditions in which there are localised or generalised abnormalities of development affecting more than one bodily system, one of which may be the skeletal system. This group of medical problems consists of more than 3000 separate conditions, most of which are genetically determined. Approximately one half has some involvement of the skeletal system. The number of syndromes increases each year in response to increasing knowledge, especially promoted by advances in clinical and molecular genetics.

Dysostoses may be defined as skeletal malformations occurring singly or in combination. The dysostoses are static and their malformations occur during blastogenesis (the first eight weeks of embryonic life). Those in which the underlying genetic mechanism has been identified have also been included in the International Classification [1].

For the purpose of establishing a diagnosis from the clinical and radiological findings, it is necessary to consider these three groups as a single entity because all present with skeletal findings. Even with increasing knowledge of specific gene mutations resulting in individual conditions, initial clinical and radiological examination is essential for diagnosis and allows targeted genetic testing to confirm a diagnosis.

Although individually rare, collectively these conditions represent a frequent problem affecting approximately 1% of the population. They are a major cause of stillbirth and disability in children. In both human and economic terms the costs are high. Individual skeletal dysplasias are rare, but overall there is a birth prevalence of 240 / 320 per million (0.3%). This figure includes those which are lethal [2,3]. An estimate of the prevalence of skeletal dysplasias requiring orthopaedic management has been made

from the dedicated skeletal dysplasia orthopaedic clinics in England and Scotland. Approximately 10,000 affected individual were identified and this did not include any paediatric, neonatal or obstetric centres, or those conditions which are mild and rarely present for treatment such as hypochondroplasia and dyschondrosteosis [4]. Accurate diagnosis requires a multidisciplinary approach, but it is heavily dependent on the identification and interpretation of sometimes subtle radiographic abnormalities.

Early and accurate diagnosis is vital if appropriate and timely advice, in the form of genetic counselling, and management and treatment for a particular disorder is to be given. Only when the diagnosis is known, can an accurate prediction of any disability be made, from understanding and knowledge of the natural history of the disorder. The effects of serious complications such as heart defects, blindness and deafness, can often be prevented or reduced by timely intervention.

Neurological complications as a result of instability in the cervical spine causing cord compression, can be prevented by spinal fusion, for example in Morquio disease. In this condition it is the combination of absence of the odontoid peg together with marked ligamentous laxity, which results in the instability at the level of C1-C2 [5]. In achondroplasia thoraco-lumbar cord compression resulting from spinal stenosis and progressive kyphosis can be prevented. The spinal stenosis is caused by a combination of narrow interpedicular distances in the lumbar spine, together with short pedicles. Many patients with achondroplasia also develop hydrocephalus as a result of a small foramen magnum and a small odontoid peg with some instability may also contribute to cervical cord compression [6]. These potential complications need active investigation and prevention. In some dysplasias cranial nerve compression is a recognised complication and in severe osteopetrosis decortication of the optic foramina prevents the onset or progression of blindness [7].

Occasionally, curative treatment of the disorder is possible as, for example, in the severe, neonatal, autosomal recessive form of osteopetrosis, by compatible bone marrow transplantation [8].

In some conditions surgical techniques enable limbs to be lengthened. This has been widely used in patients with achondroplasia in whom redundant soft tissues reduce some problems associated with bone lengthening. Although upper limb lengthening is

relatively uncommon, it is sometimes employed in patients with Madelung deformities (dyschondrosteosis) or reverse Madelung deformities (diaphyseal aclasis and Ollier's disease) to prevent disability from progressive dislocation of the radial heads. Timely orthopaedic long bone osteotomies can prevent progression of, and correct, limb deformities such as coxa vara and valgum and genu varum and valgum. Progression of hip dysplasia to dislocation can be prevented. This is a recognised feature in the natural history of many dysplasias, for example the mucopolysaccharidoses. Craniofacial deformities may be corrected especially in the craniosynostosis group of conditions and also in some of the craniotubular disorders and in fibrous dysplasia, which are associated with facial overgrowth.

Well-timed growth hormone treatment may also make an important difference to the final height of an affected individual.

The correct management of a patient relies on establishing a diagnosis, undertaking appropriate investigations and instituting curative, preventative, corrective or cosmetic procedures aimed at improving the quality of life of the affected individual.

Only when the accurate diagnosis of an individual condition has been made, can meaningful genetic counselling be given to the patient or to the parents.

Diagnosis depends on a multidisciplinary approach. However it is the correct identification and interpretation of radiological findings on a radiographic skeletal survey, which is of paramount importance in the initial evaluation process. Radiologists are trained to develop a systematic approach to the evaluation of a skeletal survey and examine all parts of the body for, firstly major abnormalities and then more subtle changes. Interpretation depends on knowledge of normal anatomy and normal variants. But it is the normal development and growth through fetal life, infancy and childhood that is particularly relevant in the domain of skeletal dysplasias. The specific, diagnostic radiographic changes are always apparent before adolescence. This is the field of radiologists specialising in paediatric radiology and is sometimes specific knowledge, which may not be readily available to general radiologists. Correct identification of the radiographic abnormalities is essential in attempting a diagnosis and currently all aids to diagnosis rely on this input. Future developments will include computerised pattern /

shape recognition from radiographic images, helping to reduce the human error of incorrect feature identification.

The standard method of diagnosis of skeletal dysplasias by general radiologists is by identifying abnormalities and referring to standard textbooks [9,10,11,12] or cross-referencing lists of gamuts or handles [11,13,14]. Images from the skeletal survey, when possible diagnoses have been identified, are compared to textbook images if these are available. Unfortunately, because of size constraints, adequate numbers of images, covering different ages, cannot be available in standard textbooks. In addition the text describing the dysplasia may not provide a visual description of the images because of different terminology or limited descriptions. Also, each condition will show some heterogeneity with a range of findings in addition to the changes expected to occur with time, adding to the difficulty of making a true match. General radiologists will also consult with other clinical colleagues and finally refer to an expert in the field. Unfortunately there are only relatively few paediatric radiologists in the UK and only a handful of these have interest or experience in the diagnosis of skeletal dysplasias.

Because of the difficulties experienced by general radiologists in diagnosis in this area, computerised aids to diagnosis have been developed, firstly a knowledge-based expert system and secondly an interactive image database.

The aim of an intelligent (knowledge-based) system is to arrive at a diagnosis (or group of close differential diagnoses). An expert knowledge-based system for diagnosis captures the expertise of the diagnostic strategy or approach of experts to provide diagnostic pathways and validation processes using an interactive approach with the user through directed questions. It is designed to be used by non-experts in the field of genetic bone disorders, although by competent clinicians (in the examination of radiographs). Several expert systems have been developed in other medical domains, but this is the first in the field of skeletal dysplasias. An expert system differs from an electronic database in that it incorporates the background knowledge, used by experts, to make use of the foreground knowledge – the features of dysplasias and malformation syndromes – and incorporates diagnostic reasoning pathways. In this domain, background knowledge includes detailed knowledge of the skeletal system, its development and maturation and knowledge of normal variants. It includes common-sense reasoning and the ability to make inferences. The expert system still relies on the

input of clinical and radiological features identified by radiologists and needs to be sufficiently robust to allow for some errors of observation.

The second computerised system to be developed was an interactive database of images (a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias, REAMS) [15]. Databases, when used as a diagnostic aid, are for use by clinicians with some experience in the field - able to use their own diagnostic experience – using the database as a computerised cross-referencing tool. REAMS relies on a matching process of the findings associated with each image. It aims to incorporate images showing the full range of findings in a condition at different ages and can be more inclusive than textbook descriptions. The final diagnosis is achieved by the clinician from examining the features of the diagnoses suggested from the cross-referencing process. An important additional role of databases is that they are available for browsing as a teaching / learning facility in much the same way as an enhanced textbook may be used. Because of this improvement and the extended role and performance of databases compared to textbooks, electronic publishing companies have marketed them. On the other hand, expert systems without the direct learning / teaching objective, aimed solely at achieving improved diagnostic accuracy, are currently not being published for fear of liability in the event of an inaccurate diagnosis being made. This applies particularly where the system is not totally computerised and the input (radiological observations) is still subject to observer error.

It is anticipated that some of the intelligent functions of an expert system could be incorporated into a database to further enhance diagnostic accuracy, without increased liability on the part of the publishers.

Chapter 2. Background

2.1 Retrospective review of skeletal surveys referred for a tertiary opinion in the field of skeletal dysplasias and malformation syndromes during a one-year period.

The retrospective review aimed to determine the accuracy of diagnoses by the referring clinicians. This was to determine the scale of the problem of misdiagnosis with the consequent impact on patient advice and management and overall patient outcome. It cannot evaluate the problem of non-referral as a result of incorrect diagnosis and false confidence.

Unfortunately in any assessment of this kind, establishing a gold standard is extremely difficult. As a radiological expert in this field I have used my own opinion together with available clinical and genetic information in consultation with clinical colleagues, as the gold standard. I acknowledge that this may be relatively flawed. However, recognising that the figures may not stand up to scientific scrutiny, I believe the findings confirm the impression of poor diagnostic accuracy. Comparison of diagnostic accuracy between clinical specialities is likely to be valid.

In the field of skeletal dysplasias, originally the gold standard relied on quite subjective evaluation of clinical and radiological observations. However more recently biochemical studies have helped to confirm diagnoses, as in the mucopolysaccharidoses. With dramatic improvement in identification of specific gene mutations for individual dysplasias there has been confirmation of correlation between clinical and radiological phenotype and genotype. Modern mapping and detection methods have helped to validate purely clinical/radiological diagnoses.
The International Nosology and Classification of Constitutional disorders of Bone (2001) [1] lists the skeletal dysplasias currently identified and their gene mutations and pathogenesis when known. In spite of our dramatic improvement in knowledge of mutations in individual conditions, clinical and radiological evaluation of an individual patient will remain the foundation for postulating a diagnosis, for later confirmation by molecular genetic testing. However in some conditions the identification of the specific mutation may not contribute to establishing the precise diagnosis. For example, in type II collagen mutations, the clinical phenotype usually cannot be predicted from the individual genetic mutation, and the diagnosis may range in severity from perinatally

lethal achondrogenesis type II or hypochondrogenesis, to SEDC with severe short stature, or to Stickler syndrome with normal stature and life expectancy. It is presumed that there are modifiers of the primary mutation (as yet unknown) to explain this lack of clinical correlation.

Another factor influencing the gold standard will be the literature description of further cases of rare disorders, expanding the phenotype and refining diagnostic criteria.

The definition of the gold standard is thus constantly changing in line with improved precision in clinical and radiological observations and correlation with biochemical and molecular genetic understanding and testing.

2.2 Materials and Method

307 cases were referred for my opinion during the one-year period 1998-1999. They were from around the country and did not include my main referral base of patients of Great Ormond Street Children's Hospital. They were referred for one of three reasons.

1. To confirm a suggested diagnosis.
2. To confirm that the diagnosis is unknown.
3. To establish a diagnosis when one has not been suggested.

In none of these cases was there a firm diagnosis before referral. All of the cases will have had a radiological opinion and most will have been seen by several specialities. By definition the referred cases did not include more common diagnoses where the diagnosis was certain and did not need confirmation, nor those where a certain, but incorrect, diagnosis had been made. These latter would represent an inappropriate level of false confidence.

The cases were evaluated for the following information

1. Referral diagnoses – **suggested** or **unknown**

2. Accuracy of, or agreement with, referral diagnosis. An **accurate** diagnosis was -

   Suggested diagnosis **confirmed**

   Unknown referral remained **unknown**

   An **inaccurate** diagnosis was –

   Suggested diagnosis was not agreed and **changed** to a new diagnosis

   Suggested diagnosis was not agreed and changed to **unknown**

   Unknown referral was **diagnosed**

3. An estimate of patient outcome. This was divided into –

   Optimal patient outcome –

   Suggested diagnosis **confirmed**

   Suggested diagnosis **changed**

   Unknown diagnosis **diagnosed**

   The optimal patient outcome occurred when a firm diagnosis could be established thus enabling correct clinical management and genetic counselling to be given.

   Improved patient outcome –

   Suggested diagnosis **unknown**

The improvement in patient outcome in this situation means that there is no inappropriate 'labelling' with consequent inaccurate genetic advice or clinical management.

Unchanged patient outcome

Unknown diagnosis still **unknown**

In this situation the natural history of the disorder remains unknown and firm genetic counselling cannot be given. The situation is unchanged from before the referral was made.

4. The range of diagnoses made by me.

5. The type of clinician referring the case and accuracy within these groups.
   These were identified as –
   - Geneticists
   - Paediatricians
   - Radiologists
   - Histopathologists
   - Orthopaedic surgeons

6. The major referral centres and the accuracy, or agreement with, the individual referral centres.
   These were
   - Guy's Hospital
   - Leeds
   - Institute of Child Health (non-GOSH patients)
   - Kennedy-Galton Centre
   - Nottingham

   An assessment of accuracy was -
   - **Suggested** diagnosis **confirmed**
   - **Unknown** referral remained **unknown**

7. The age of the patient. The cases were divided into four age ranges – fetus (mainly elective terminations of pregnancy at about 21 weeks gestation), 0-10 years, 11-16 years and over 16 years and the number of cases in each age range identified. In the different age ranges an evaluation was made of the percentage

referred with a suggested diagnosis, the accuracy of the suggested diagnosis from either a confirmation of the suggested diagnosis, or agreement that an unknown referral remained unknown, and cases in which a diagnosis was confirmed. This was those cases where a suggested diagnosis was confirmed, a suggested diagnosis was changed and an unknown diagnosis was changed. This was a representation of optimal patient outcome.

2.3 Results

Referral diagnoses – suggested or unknown

Table I

Referral diagnoses

suggested referral diagnosis	146	48%
unknown referral diagnosis	161	52%
total referrals	307	100%

suggested diagnosis **confirmed**	67	46%
suggested diagnosis **changed**	43	29%
suggested diagnosis **unknown**	36	25%
total with **suggested** diagnoses	146	100%

unknown diagnosis **changed**	85	53%
unknown diagnosis remained **unknown**	76	47%
total with **unknown** diagnosis	161	100%

Results from the whole group

suggested diagnosis **confirmed**	67	22%
suggested diagnosis **changed**	43	14%
suggested diagnosis **unknown**	36	12%
unknown diagnosis **changed**	85	28%
unknown diagnosis remained **unknown**	76	25%
total referrals	307	100%

Of the referred cases, about half had a suggested diagnosis and about half were unknown. Of the cases with a suggested diagnosis this was confirmed in about half of them. In about a third of cases (37%) the final diagnosis was unknown. This was where a suggested diagnosis was changed to unknown and where the referral, unknown diagnosis remained unknown. Conversely, a diagnosis was established in two-thirds of cases.

Accuracy of diagnosis

Table II

Accurate diagnosis (referral diagnosis in agreement with my opinion)

suggested diagnosis **confirmed**	67	22%
unknown diagnosis remains **unknown**	76	25%
total with an **accurate** diagnosis	**143**	**47%**

Inaccurate diagnosis (referral diagnosis not in agreement with my opinion)

suggested diagnosis **changed**	43	14%
suggested diagnosis **unknown**	36	12%
unknown diagnosis **changed**	85	28%
total with an **inaccurate** diagnosis	**164**	**54%**

About half the cases referred had an accurate referral diagnosis, and half were inaccurate.

Patient outcome

Table III

optimal	**suggested** diagnosis **confirmed**	67	22%
	suggested diagnosis **changed**	43	14%
	unknown diagnosis **changed**	85	28%
	total	**195**	**64%**
improved	**suggested** diagnosis **unknown**	**36**	**12%**
unchanged	**unknown** diagnosis remains **unknown**	**76**	**25%**

About two-thirds of patients achieved an optimal outcome with a firm diagnosis being established allowing planned management and correct genetic counselling. In a quarter of cases there was no improvement in patient outcome.

Range of diagnoses

Table IV

Common diagnoses.

Normal	47
Osteogenesis Imperfecta	13
Rickets	7
Asphyxiating Thoracic Dystrophy	6
Multiple Epiphyseal Dysplasia	5
X-linked rickets	4
Thanatophoric dysplasia	4
Achondrogenesis	3
Acrodysostosis	3
Chondrodysplasia punctata	3
Cleidocranial dysplasia	3

Rare diagnoses. There was only one case of each of these conditions.

Atelosteogenesis type I

Desbuquois Dysplasia

Femoral Facial Syndrome

Greig Polysyndactyly Syndrome

Hand Foot Genital Syndrome

Hypochondrogenesis

Kyphomelic Dysplasia

Microcephalic Osteodysplastic Primordial Dwarfism

Neu Laxova Syndrome

Robinow Mesomelic Dysplasia

Short Rib Syndrome type Beemer-Langer

Although these were the most common cases referred, they do not represent the most common conditions. Osteogenesis imperfecta is common, but the majority of these referred cases were fetuses and the bent/angulated bones resulted in diagnostic difficulties. Thanatophoric dysplasia is equally common, but is more readily identified on prenatal ultrasound and clinically, following termination of pregnancy. The high proportion of cases of rickets was surprising and was misdiagnosed by radiologists and by paediatricians because of normal biochemistry, with the rickets in the process of healing. Achondroplasia is the most commonly encountered

surviving dysplasia and is notably absent from these referrals because it is readily diagnosed both clinically and radiologically by non-experts.

Accuracy by clinical specialty

Table V

Clinical specialty	number of cases out of 307	% of total cases	% accuracy
Genetics	152	50%	57%
Paediatrics	86	28%	36%
Radiology	36	12%	31%
Histopathology	29	9%	34%
Orthopaedics	4	1%	sample too small

Clinical geneticists referred half the cases and they had the highest accuracy rate of 57%, when compared to the overall accuracy rate of 46%. Radiologists achieved the lowest accuracy of 31%. There are several potential reasons for this difference. Geneticists devote their professional time to this precise domain of genetically determined conditions. Because of the individual rarity and extensive range of conditions geneticists make use of databases and computer assisted diagnosis in their day-to-day work [16,17,18,19,20] and are more familiar with the range of possibilities. However they are not able to evaluate fully the radiological findings, probably explaining the high referral rate.

The reasons for this poor result by radiologists are undoubtedly multifactorial. They relate to the limited time in training, the breadth of the curriculum, which includes practical experience of multiple imaging modalities, a knowledge of adult and paediatric pathology, normal variants and normal paediatric development and also varied practical and emergency procedures. The domain of skeletal dysplasias and malformation syndromes is large and includes many rare disorders with varying combinations of skeletal findings. Radiology training does include a requirement for some knowledge of the more common skeletal dysplasias and plain radiographic and skeletal evaluation. The final Fellowship examination of the Royal College of Radiologists (FRCR) in the UK reflects this. Inevitably the full range of conditions cannot be encompassed during training and there is a reluctance to further specialise in this field after radiological accreditation.

Accuracy of main referral centres

Table VI

Referring Centre	number of cases	% accuracy
Guy's Hospital	45	60%
Leeds	20	30%
Institute of Child Health	15	66%
Kennedy-Galton Centre	11	82%
Nottingham	9	22%

These constitute about one third of the total number of cases. The remaining referrals from other centres were of fewer numbers. Probably little can be inferred from these findings. Although higher accuracy was achieved by London centres, this may have been because they were more willing to refer more cases, and more straightforward cases.

Breakdown by patient age

Table VII

	Fetus	0-10 years	11-16 years	over 16 years	All
Number of cases	49	176	41	38	307
Referral diagnosis suggested	57%	41%	61%	58%	48%
Accuracy	39%	47%	49%	53%	46%
Diagnosis confirmed or changed	88%	66%	59%	61%	68%

As would be expected, most cases present in the first 10 years of life. Accuracy was about 50% in all postnatal groups, but was much lower in the fetal group (39%). In the fetal group a diagnosis could be established in 88%, which was much higher than in the other age ranges. This may be because some were referred with a pre-termination of pregnancy ultrasound diagnosis, which can only be a general guide to severity of the condition and often cannot be precise. Prenatal ultrasound reliably predicts severe skeletal dysplasias with a poor outcome, but can suggest an accurate diagnosis in only about 50% of cases. In addition from the post-termination skeletal survey there is less familiarity by radiologists in general with normal radiological findings at different gestational ages, and also what findings would be expected at 20 / 22 weeks gestation following termination of pregnancy, of dysplasias normally presenting at birth. This means that the fetal cases were probably less selected in their referral pattern.

2.4 Conclusions and Discussion

The retrospective review highlights the problem of accurate diagnosis in the field of skeletal dysplasias and malformation syndromes. Diagnosis depends to a large extent on the accurate evaluation and interpretation of radiographic skeletal findings. Whilst radiologists are trained in the examination of radiographs and are by definition diagnosticians, they perform poorly in this particular area, achieving a diagnostic accuracy of 31%, lower than other clinical groups with no particular skills in radiological interpretation (average accuracy 46%).

The current diagnostic approach among radiologists is to refer to large standard textbooks and atlases of skeletal dysplasias [10,11,12], to consult with colleagues or an expert or to refer to lists of gamuts using cross-referencing, sometimes referred to as triangulation [11,13,14]. There are time constraints and often lack of motivation. Currently in the UK there are only a handful of paediatric radiologists with experience in the diagnosis of skeletal dysplasias. This field lends itself to computer-assisted diagnosis because of the breadth of the domain, the lack of expertise and the positive impact on individual patients and their parents of an accurate diagnosis allowing accurate genetic counselling and appropriate management and sometimes treatment.

This retrospective review provides some background information confirming the clinical impression that clinicians in general and radiologists in particular, achieve poor diagnostic accuracy in the UK. It has provided the impetus to develop two computer based systems aimed at improving diagnostic accuracy in the field of skeletal dysplasias -

- A knowledge-based expert system
- An image database

Chapter 3. Development of the expert system

The expert system aims to model the diagnostic skills of two radiologists, expert in the diagnosis of skeletal dysplasias (myself and DS). The skills were elicited through extensive consultation interviews between the knowledge engineers (EK and JW) and the radiologists, independently and collectively and refined in parallel with the development of the knowledge base and through trials of the system. The diagnostic reasoning model was designed to represent the reasoning of the experts and as such provides a dynamic system [21]. Several computer-based diagnostic aids for the domain have been reported [9,13,16,17,18,19, 20, 22, 23, 24, 25, 26]. Some are conventional database systems. The domain coverage of these systems varies and none aims to model the skills of experts. The combination of explicit representation of background knowledge with a specifically designed diagnostic reasoning model is what distinguishes a knowledge-based system from a database system in which information is searched using standard database search facilities [21, 27, 28, 29, 30, 31, 32].

The expert system comprises:

a diagnostic reasoning model

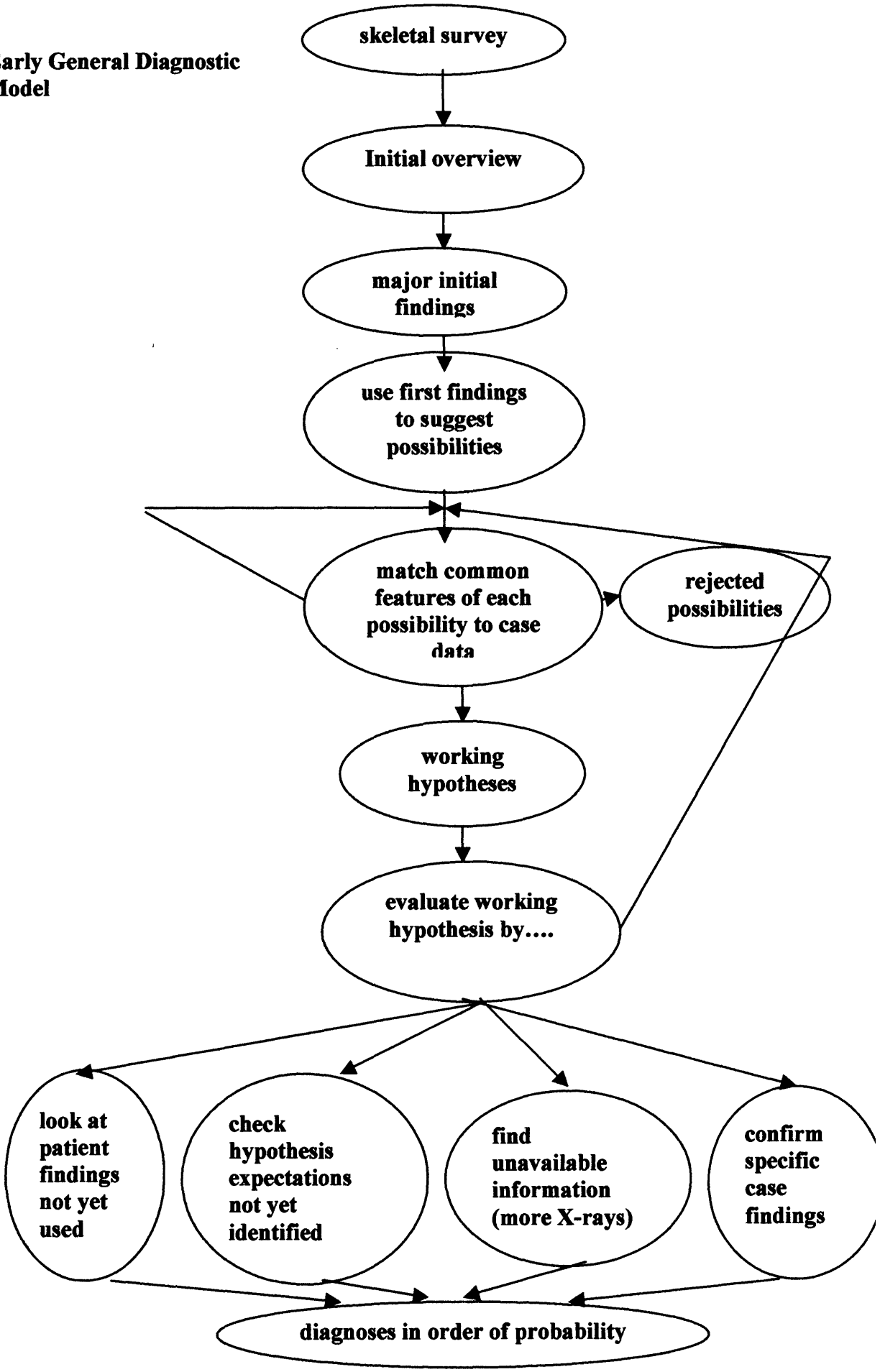
foreground knowledge about individual dysplasias

background knowledge relating to the skeletal system

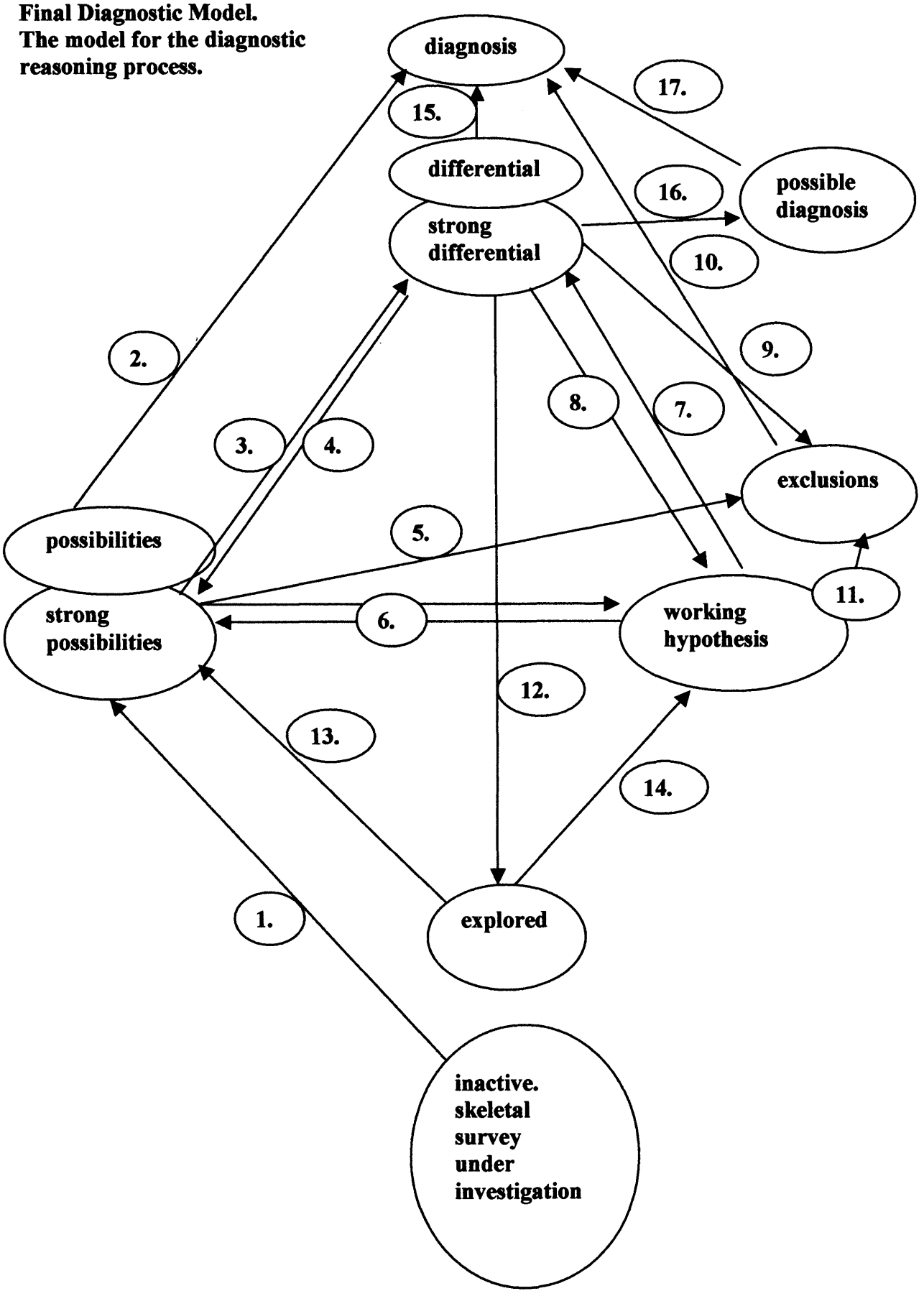
3.1 Development of the diagnostic model

After initial sessions using individual case models and discussing the diagnosis of dysplasias individually by the two experts, and also collectively, a preliminary diagnostic model was established. This was deliberately general, aiming to explore the early stages of development of a diagnostic model without arriving at premature conclusions or perpetuating misrepresentations. Refinements, following further discussions and testing rapidly followed, leading to the final diagnostic model [21].

Early General Diagnostic Model



Final Diagnostic Model.
The model for the diagnostic reasoning process.



Pathways for the Diagnostic Model

1. Primary or secondary trigger match.
2. Secondary trigger match
3. There are no working hypotheses and the possible hypothesis (possibility) has a higher score and is promoted.
4. The possible hypothesis no longer has a high score from the explanatory power and sufficient set match.
5. Exclusion criteria are met.
6. Possible hypothesis promoted to working hypothesis if it has a sufficiently high hard abnormalities coverage; or a working hypothesis is demoted to a possible hypothesis if the coverage is inadequate.
7. Working hypothesis has a relatively high sufficient set match.
8. Working hypothesis no longer has a relatively high sufficient set match.
9. Exclusion criteria satisfied.
10. Typical feature match.
11. Exclusion criteria satisfied.
12. Hypothesis completely explored but no firm conclusion can be reached due to incomplete information.

13,14. Relevant missing information becomes available.

15. Hypothesis strong enough and sufficiently better than its closest competitor and/or suggested categorically by closest active competitors; hypothesis also has a sufficiently high hard abnormalities coverage.

16. Missing information prevents the differentiation of the strong differential.

17. Missing information becomes available.

The specific terms used in the Diagnostic Model are explained in the subsequent text.

3.2 Foreground knowledge and components of the dysplasia frame

After the lengthy process of exploring the reasoning involved by two experts in dysplasias (myself and DS) in making a diagnosis and in the various pathways followed, the knowledge engineers (EK and JW) created a preliminary diagnostic model, which has evolved into the final diagnostic reasoning model as above [21]. The expert data input model was defined in the form of the **dysplasia frame,** which contained all the foreground knowledge about the individual dysplasias.
I had the responsibility of entering the data into the dysplasia frames, and created them for approximately 200 conditions, creating about 2000 distinct features.

Foreground knowledge about dysplasias is the features identified and entered by experts. Features of a dysplasia together describe the dysplasia. Each feature is a concise description and consists of a subject (a precise anatomic part) and various attributes.

Features are differentiated according to their diagnostic significance. They include **common** features, which have been observed to occur in the majority of cases. Given that a patient suffers from this particular dysplasia the **common** features would be expected to be present. **Occasional** features (represented by '+-' in Appendix A) are abnormalities which would not be expected in every case. Their presence counts in favour of the hypothesis of the dysplasia, but their absence does not count against.

Each feature is also identified as **hard** or **soft**. **Hard** abnormalities are very significant from a diagnostic point of view, but **soft** abnormalities are relatively non-specific.

Each feature of a particular condition is broken down into the 'speech parts' of that term. There are 10 principal speech parts and each word belongs to only one of these. These are shown in **Table VIII.**

Table VIII

Foreground knowledge and components of the dysplasia frame

Table of Principal Speech Parts

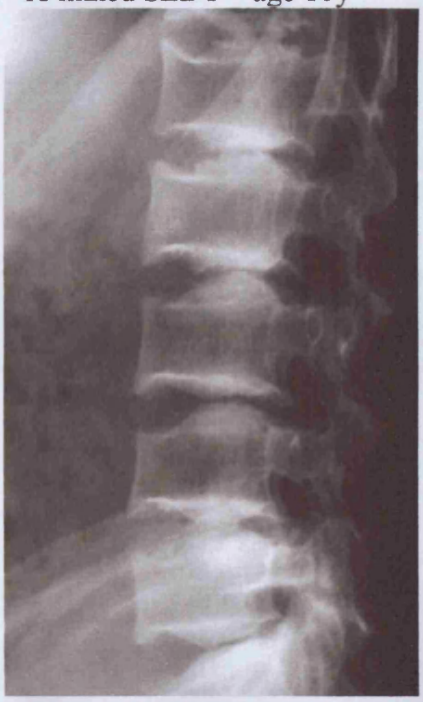
Speech Part		Examples
Subjects	Anatomy Body Part System Histology Biochemistry	skull, liver head, abdomen skeletal system erythrocytes, epithelial cells urea, keratin sulphate
Quantifiers		all, some, few
Descriptors		long, curved
Qualifiers	degree distribution relative to	severely medially, scattered tibia relative to the fibula
Processes		ossification, maturation
Conjunctions		and, not
Prepositions	general relations	with
	spatial relations	above, behind
Temporal terms		from, to, at
Temporal values		age and age ranges, neonate
Temporal qualifiers		progressive, increasing

Features

The features are grouped into 4 sets. These are listed in decreasing order of diagnostic weighting.

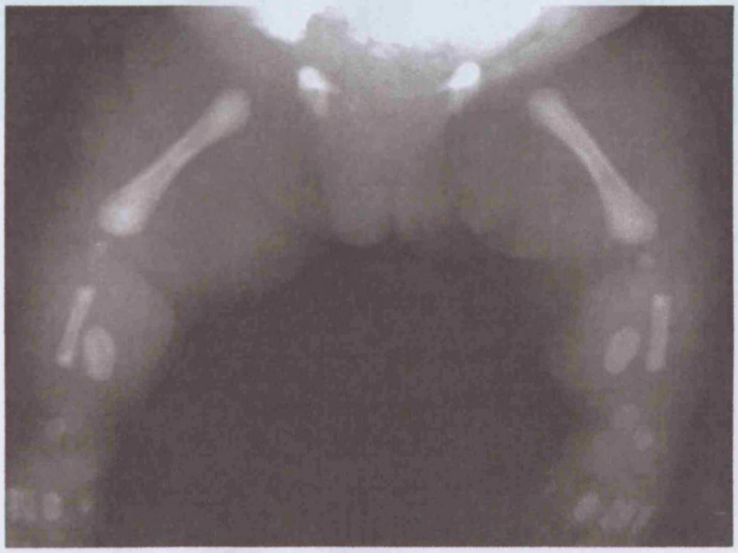
1. **Typical** features are rare but conclusive of a diagnosis. The absence of the feature should not count against the possibility of the dysplasia. One example is the typical appearance of the vertebral bodies in the older child with X-linked spondyloepiphyseal dysplasia tarda (SEDT) in which there are dense mounds of bone on the posterior two-thirds of the vertebral end plates as seen in **Figure 1.**

Figure 1 X-linked SEDT – age 10y



Another example of a single feature forming a **typical feature** is the short, oval tibiae in Short-Rib-Polydactyly Syndrome Type II (Majewski) in **Figure 2.**

Figure 2 Short rib polydactyly syndrome type II - stillbirth



More commonly a group of a few features together become diagnostic of a condition.

For example, in Dyggve-Melchior-Clausen disease (DMC), the iliac crests are irregular and 'lace-like'. This finding is also seen in adenosine deaminase (ADA) deficiency, some cases of Ollier disease and rarely in unusual metaphyseal chondrodysplasias. The combined features of irregular iliac crests and platyspondyly would only be seen in DMC and ADA deficiency. The three features of irregular iliac crests, platyspondyly and small irregular epiphyses would be typical of DMC as in **Figures 3a,b**

Figure 3a Dyggve-Melchior-Clausen disease showing lace-like iliac crests.

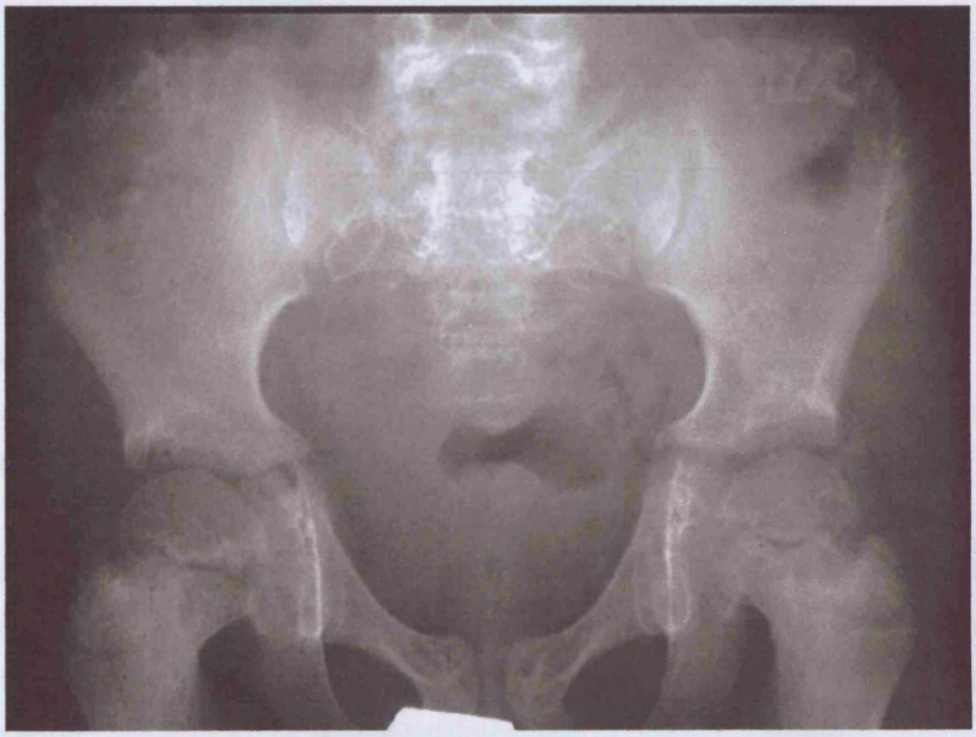
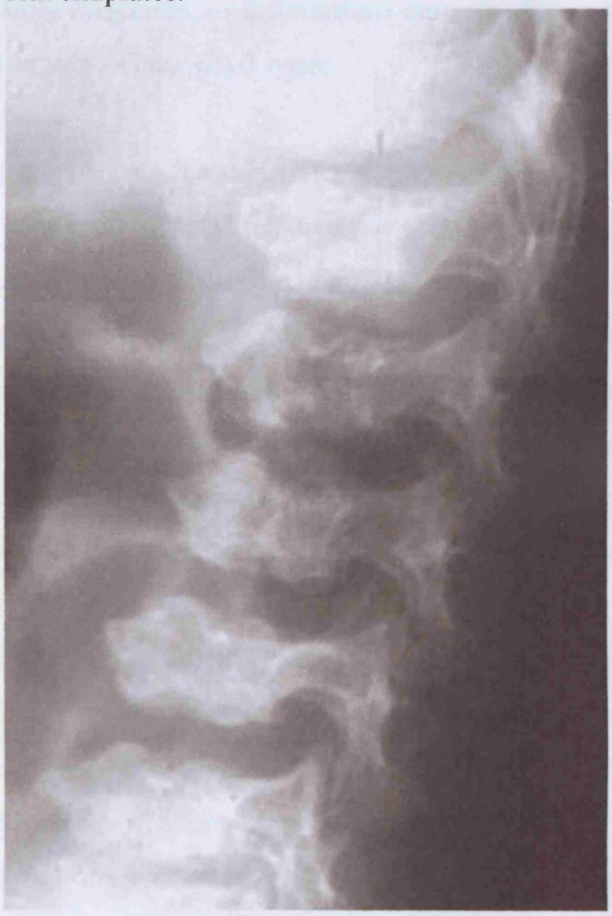


Figure 3b Dyggve-Melchior-Clausen disease showing platyspondyly and central notches on the vertebral endplates.



2. Exclude-if features, if present, are sufficient to exclude the possibility of the dysplasia. Such features are the opposite of expected features. For example the presence of 'coxa valga' would exclude a diagnosis of spondyloepiphyseal dysplasia congenita (SEDC) in which 'coxa vara' is expected.

3. Sufficient set features are a group of features, which collectively strongly indicate the possibility of the dysplasia. The **sufficient set** is made up of a group of **common** features. It does not mean sufficient to confirm as a diagnosis, but to warrant further exploration. It is possible for several **sufficient sets** to exist for a dysplasia, usually when there are several sub-types. The **sufficient set** is used at all levels of consultation and has an important role in the scoring of findings entered at the beginning of a consultation. It contributes to the order of ranking of diagnoses for consideration. The system re-evaluates the match of the **sufficient sets** at different stages and uses it to promote and demote possibilities.

4. **Triggers** direct attention economically towards possible diagnoses, or quickly eliminate impossible diagnoses, or differentiate between diagnoses with highly overlapping feature sets. There are 3 types.

    1. **Primary triggers** are features that catch the attention and are all '**hard**' features compared to '**soft**' features. They may be **common** or **occasional** features. **Hard** features are of recognised diagnostic importance, are referred to as diagnostic 'handles', and in textbooks may be represented as lists of gamuts. During the course of a consultation, any finding entered which is a **primary trigger** will automatically generate a list of possible diagnoses for consideration. In the diagnostic reasoning process, the **primary triggers** are an early trawl to encompass the range of conditions for further evaluation.

    2. **Negative triggers** are features which are sufficient to exclude a large subset of dysplasias considerably restricting the search parameters. For example 'absent ossification of the skull vault' is sufficient to exclude all but a handful of dysplasias from the entire domain. A **negative trigger** refers to a set of dysplasias which should be considered. Any dysplasias not in that set should not be considered. Thus, like **primary triggers** they bring hypotheses into consideration, but unlike them they also exclude hypotheses.

    3. **Secondary triggers** (differential diagnoses) differentiate between clusters of dysplasias which have highly overlapping feature sets. They have two functions. Firstly they constitute a secondary route (compared to **primary triggers**) to trigger possible hypotheses. Secondly they provide the means for differentiating between strong differential diagnoses. They are features that refute the diagnosis currently under consideration.

The features of each condition were entered into a **dysplasia frame** (**Tables VIII, IX and Appendix A**).

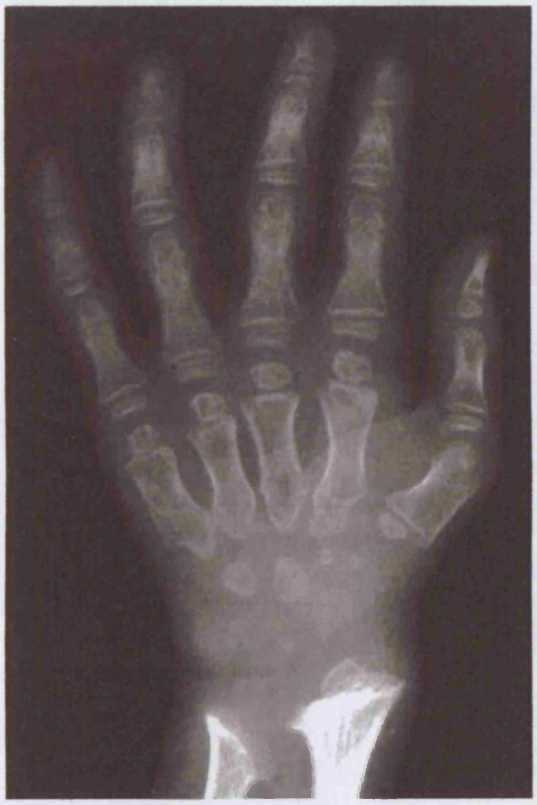
They were entered under the appropriate speech parts and grouped into categories (clinical, histological, radiological).

For example, in Morquio disease, when describing the hand, (**Figure 4a**) a feature may be that 'there is proximal pointing of the 2^{nd}-5^{th} metacarpals from the age of 2 years.' In this case the subject is 'metacarpals', quantifiers are '2^{nd}-5^{th}', the descriptor is 'pointed', the preposition (spatial) is 'proximally' and the temporal term is 'from 2 years'.

Example of a feature entry

Subject	Quantifiers	Descriptors	Preposition	Temporal term
metacarpals	**2^{nd}-5^{th}**	**pointed**	**proximally**	**from 2 years**

Figure 4a Morquio disease –age 10y- showing proximal pointing of the 2^{nd} – 5^{th} metacarpals



Each feature was identified as being **common** or **occasional** and **hard** or **soft**.

From the total list, appropriate features were selected to fulfil the categories of **typical features,** (common or occasional, hard features) if possible, **sufficient sets** (common features only, hard or soft) and **primary triggers** (common or occasional, hard features).

Close differential diagnoses of the dysplasia being entered were identified and specific, targeted features (**secondary triggers**) were entered, which pointed away from the diagnosis being entered and towards the differential diagnosis. Secondary triggers are a different form of negative trigger, only used to exclude close differential diagnoses when a specific diagnosis is under consideration.

Data entered into each dysplasia frame

Table IX

Dysplasia name	
Alternative names	synonyms
Features – clinical	features - **common** or **occasional, hard** or **soft**
- histological	entered under subject, descriptor, qualifier, quantifier etc.
- biochemical	
- radiological	
- other	e.g. prenatal diagnosis, MRI findings
Typical features	e.g. oval tibiae
Sufficient sets	selected from **common** features
Primary triggers	'handles'
Secondary triggers	features of differential diagnoses

Please see Appendix A for an example of a dysplasia frame.

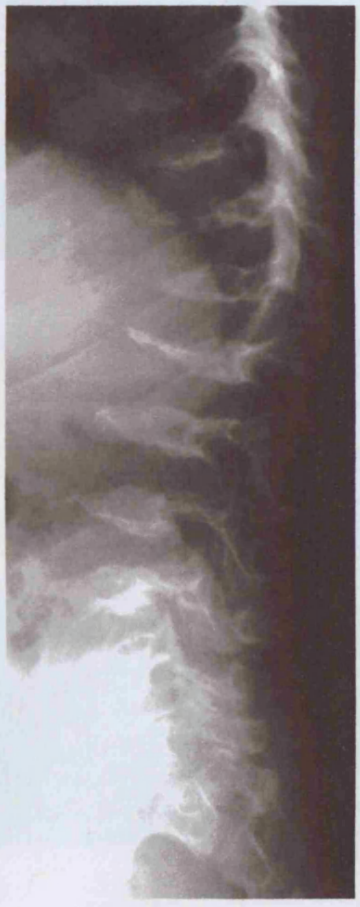
3.3 The role of secondary triggers in the differentiation process.

Morquio disease is under consideration. The close differential diagnoses are pseudoachondroplasia. spondylo-metaphyseal dysplasia, Kniest disease and metatropic dysplasia.

Case findings: - Proximal pointing of the metacarpals
Abnormal epiphyses
Coxa valga
Flared iliac wings
Sloping acetabula
progressive disappearance of the capital femoral epiphyses
platyspondyly
Figures 4a (page 37),b,c,d,e

The close differential diagnoses all have platyspondyly and therefore this feature could not be used in the differentiation process.

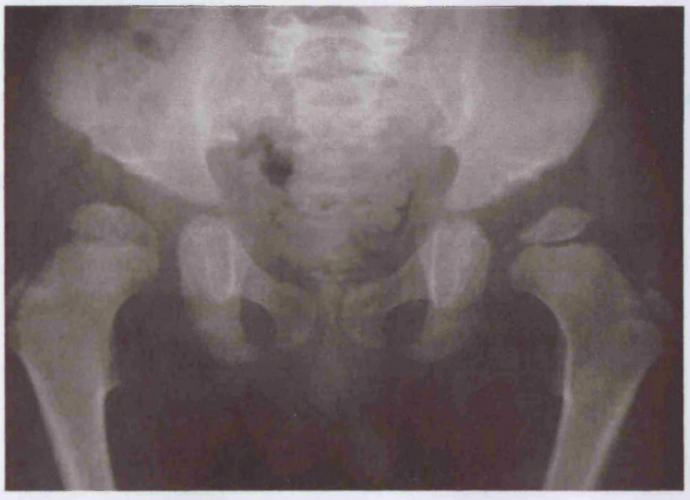
Figure 4b Morquio disease showing marked platyspondyly



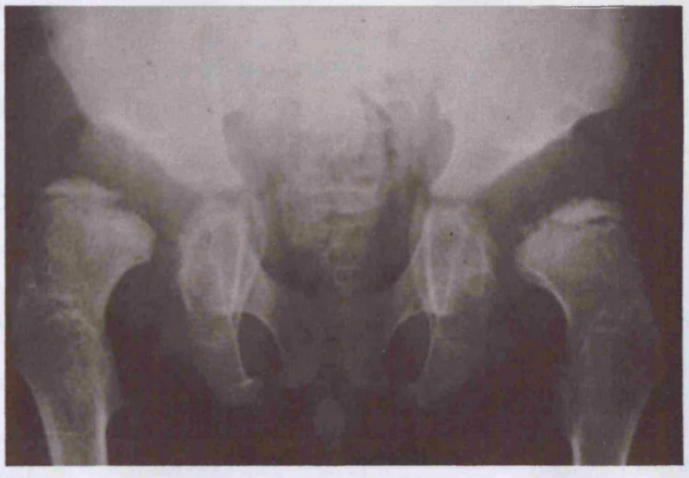
Figures 4c,d,e

Morquio disease showing progressive flattening and disappearance of the capital femoral epiphyses

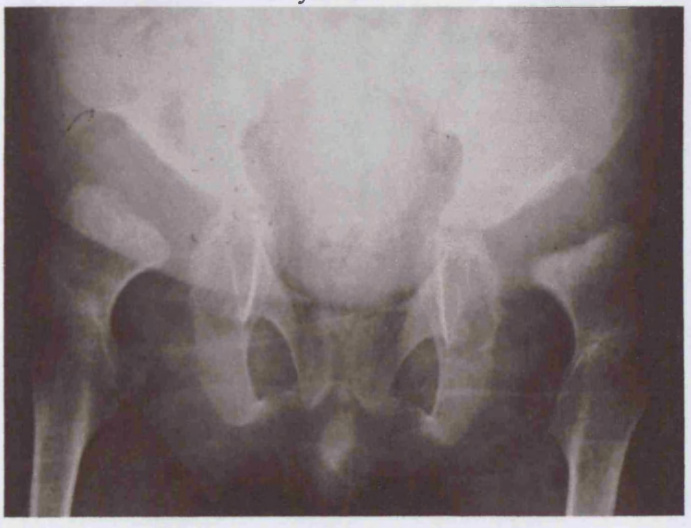
2 years 6 months



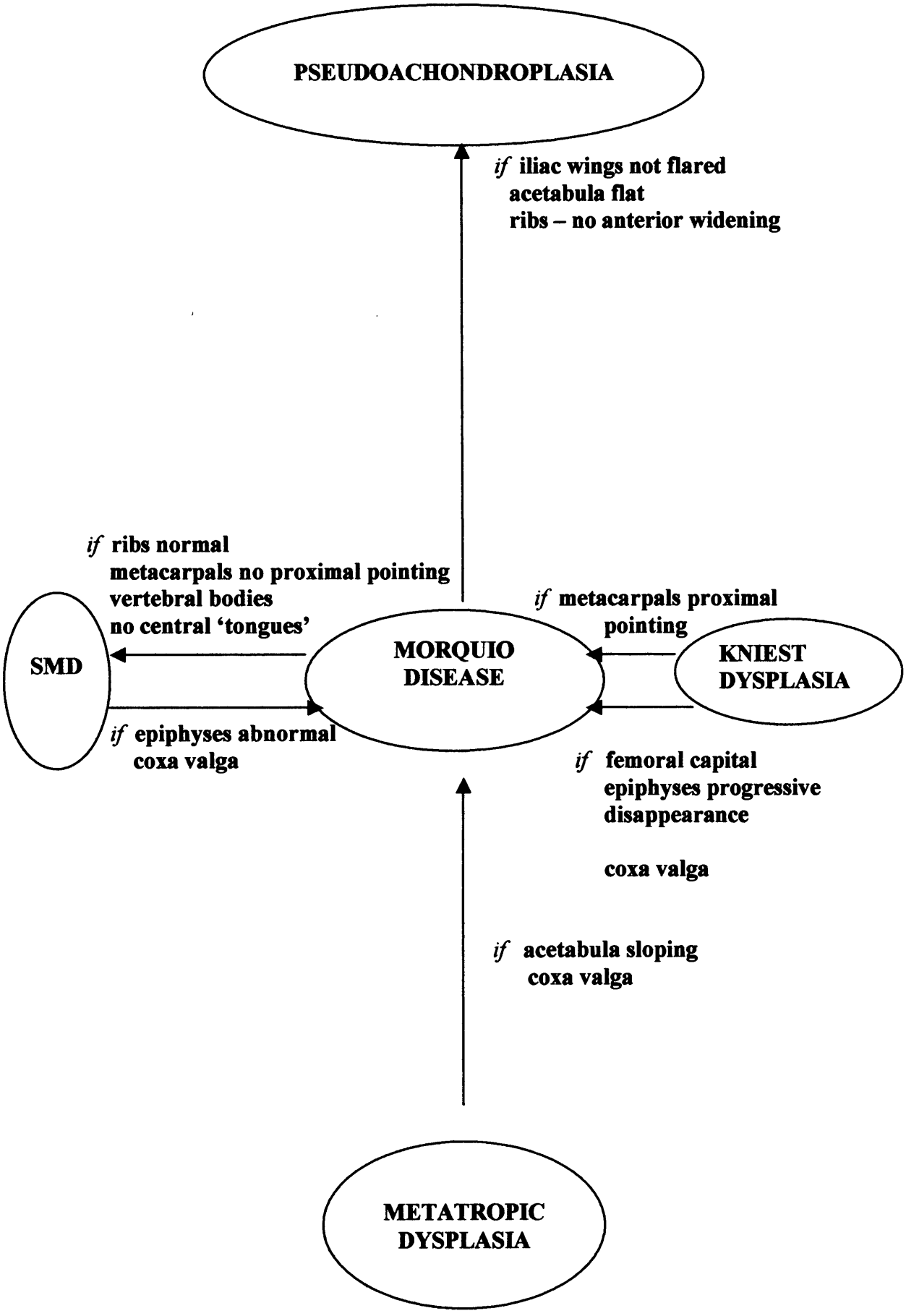
9 years



10 years 6 months



Differential diagnosis strategy (secondary triggers)

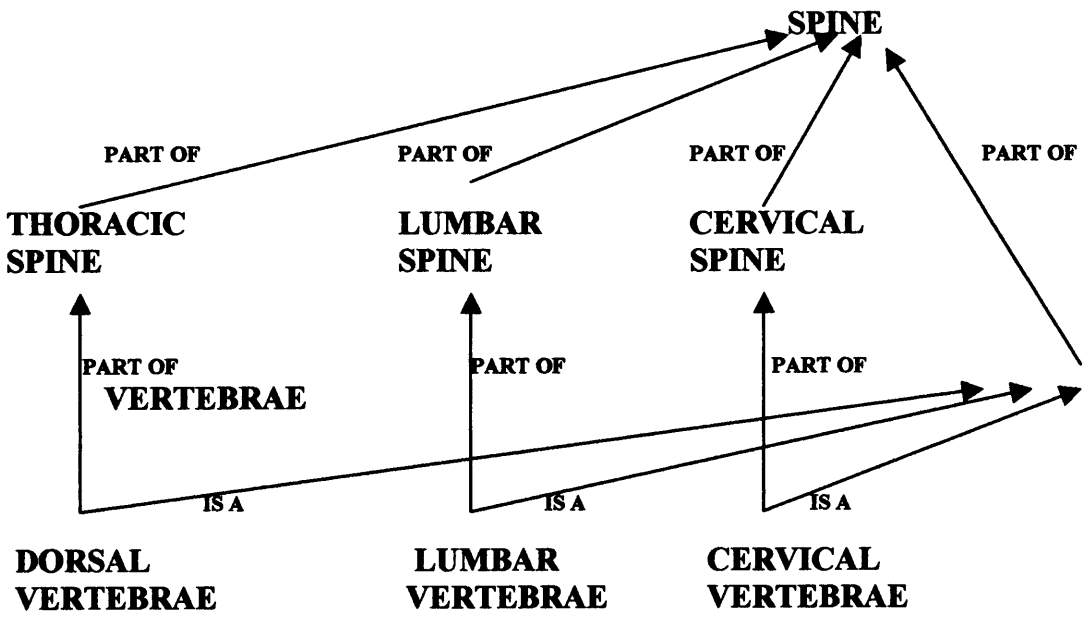


3.4 Background knowledge

Throughout the process of acquiring knowledge of individual dysplasias in the dysplasia frames I was involved in ongoing discussions with the knowledge engineers in relation to the entered features. Essentially this was to establish the background knowledge. Background knowledge is not used in the diagnostic process directly. However it is involved in the overall problem solving activity and is essential for competent behaviour of any expert system. It makes sense of case findings so that questions asked of a user of the system are intelligent [30, 31].

Background knowledge consists of **general medical knowledge** in the field of skeletal dysplasias. This includes clinical, radiological, anatomical and skeletal aspects. The system then has a deeper understanding of the features, which enables it to make intelligent inferences. For example, a user of the system may enter as a finding for an undiagnosed case 'the first lumbar vertebral body is hypoplastic'. From the **background knowledge** relating to the skeleton, the system knows that the lumbar vertebrae are part of the lumbar spine and the lumbar spine is part of the spine. Therefore it would not subsequently ask the question 'is the spine normal?'.

Processing information using background knowledge. An example of bone taxonomy.



Background knowledge includes **common-sense knowledge** about the subjects, specifically that related to the field of the radiological diagnosis of skeletal problems. It includes **taxonomic knowledge**, which identifies classes or groups such as epiphyses, metaphyses, and long bones: **meronomic knowledge**, for example if the capital femoral epiphysis is flattened, this is part of the femur and therefore the femur is abnormal. This background knowledge is an essential part of the system, enabling it to reason with information entered, thus relating findings entered by a user of the system to expected features of a dysplasia using taxonomic and meronomic reasoning [29, 30, 31, 32]. Background knowledge includes common-sense knowledge about the attributes or descriptions of subjects and includes **spatial knowledge**, for example proximal, distal, medial and lateral; and **temporal knowledge**, for example the normal progression of ossification and maturation of individual bones. Temporal aspects of reasoning will be looked at in more detail. They form important components of both foreground knowledge relating to individual dysplasias and background knowledge relating to normal growth and ossification of the skeleton.

Other important areas of background knowledge include the identification of **synonyms** and synonyms in context or synonymous subjects. This enables the flexible entry of a user's information. For example, 'wide' metaphyses and 'broad' metaphyses have the same meaning, and 'platyspondyly throughout' means the same as 'flattened vertebral bodies'.

Dependencies between findings were also explored and these enabled the system to make certain inferences. For example if there is platyspondyly throughout, then the inference is that the trunk is short. The reverse does not apply. There are many other conditions of the spine which may result in a short trunk, for example kyphosis, scoliosis or fusions.

Also, if the metaphyses are flared and the long bones have a dumb-bell appearance, then the joints are prominent or enlarged. When inferences can be made, redundant questions can be avoided during the course of a consultation. On the other hand the system needs to recognise the dependencies to avoid inappropriately high scoring from multiple overlapping findings. Similarly, **correspondence** between radiological and clinical features needs to be identified. For example, 'short radius and ulna' and 'short forearm' identify the same abnormality and again should not count twice in the scoring process.

The separation of foreground expert knowledge from background knowledge has two advantages. First, the expert knowledge is more succinct and easier both to read and maintain. Secondly, it allows the possible re-use of background knowledge in other applications, in this domain and in other medical domains.

The diagnostic reasoning employed by the experts in this field does not suggest that contending diagnostic hypotheses are ranked on the basis of single numeric estimation of their likelihood. The reasoning has identified that hypotheses are ranked from different qualitative perspectives, which are not combined into an overall ranking. The following perspectives have been identified –

- The proportion of case findings matching the common, typical or other features for the dysplasia.
- The proportion of case findings in conflict with common features of the dysplasia.
- The proportion of case findings which are irrelevant to the dysplasia.
- The proportion of the common features in conflict with case findings.
- The proportion of the common findings of the dysplasia in agreement with case findings.

These perspectives indicate how well a hypothesis accounts for the case findings and how well the hypothesis expectations are met by the case findings.

3. 5 Temporal reasoning in the diagnosis of skeletal dysplasias

Introduction

Temporal information is a vital component in diagnosing skeletal dysplasias, providing a valuable tool to help radiologists remove inappropriate conditions from consideration, and more accurately assess the degree of similarity between an individual patient and the textbook description. Any diagnostic aid should aim to incorporate explicit temporal information and therefore changes occurring with time will be presented in some detail.

The volume of data is such that current standard textbooks are unable to include all relevant temporal information and make computerised systems ideal diagnostic tools. Information about changes that occur with time is incorporated as foreground knowledge into the individual dysplasia frames as part of the entry of features in the expert system, and also in the background knowledge of the system [21, 29, 30, 31, 32, 33, 34]. Temporal changes have also been included in the image database, REAMS, a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias [15] by incorporating images of individual conditions at different ages where possible and enabling searching to be conducted by age.

Giedion in 1994 [35] identified temporal changes in genetic bone disease and stressed the importance of recognising them in arriving at a diagnosis. This he referred to as the 'weight of the fourth dimension'. Skeletal dysplasias affect primarily skeletal form and function in children and young adults. Information from radiographs and interpretation of the findings forms the basis on which a skeletal dysplasia can be diagnosed. In any one case this information will be limited to the findings present at the time the radiographs were performed. This means that changes present at an earlier age, which have subsequently disappeared, and those which have not yet appeared, cannot be identified. The interpretation of paediatric radiographs depends on the correct use of temporal information. Textbook descriptions of expected features of skeletal dysplasias, because of the volume of information, often cannot describe explicit temporal changes.

Normal development

Major changes in bone form and shape are seen during infancy and childhood, and normal processes can vary in onset and duration without being evidence of abnormality. Any evidence of abnormality therefore has to be judged against the background of

normal expectations, and the paediatric radiologist needs detailed background knowledge of the normal evolution of skeletal development in children, and their radiographic appearances. Knowledge of the range of normal variations in both morphology and maturation is also important.

There are two major difficulties faced by radiologists attempting to use temporal information in a diagnostic context. The first is that each set of radiographs represents a single time-slice, and often this is all the radiologist will have. It is difficult from a single set of radiographs to assess the age of onset of a particular abnormality, and also the stage that the abnormal process has reached, and the final result of the process.

The collective radiographic features of individual conditions have a natural history, or evolution, so that each condition may exhibit differences in utero, at birth, in infancy, through childhood to adult life.

Furthermore, aspects of the skeleton that appear normal may be the result of either normal development, or a phase within an abnormal process. For example, in the mature skeleton, after fusion of the epiphyseal plates, many diagnostic features relating to metaphyseal changes are obliterated and the metaphyses then appear normal.

Also, changes on the skeletal survey may represent consequences or complications of the primary skeletal problem. For example, abnormal epiphyseal development as part of a skeletal dysplasia will develop secondary degenerative changes superimposed on the original features. These consist of premature osteoarthritis with joint deformity and contractures, which do not directly help in diagnosing the underlying causative condition.

The second is that while textbooks present a description of expected abnormalities, they often do not include complete temporal information. Each abnormality characterising a dysplasia will have an expected age of onset and duration, with the result that the overall appearance of a dysplasia may vary quite widely over time.

This information may not be explicitly presented in textbooks, but has to be extrapolated from the descriptions given. The radiologist is required to study the textbook description and try to construct a picture of what abnormalities would be expected given the age of the patient. In addition, as many features can vary in age of

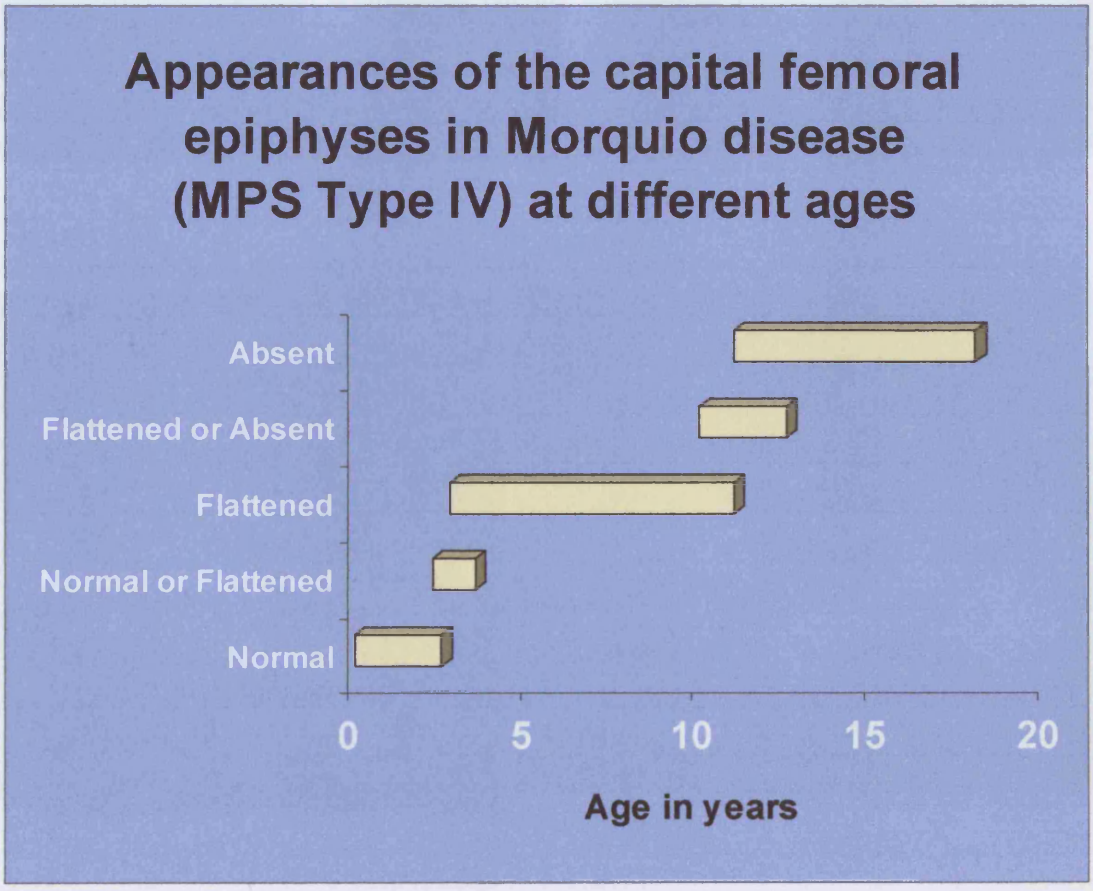
onset and duration, the radiologist has to allow for a range of different possible presentations.

For example, in Morquio disease, the changes in the appearances of the capital femoral epiphyses evolve through two critical periods: **Table X**

1. At about age 2y there is a change from normal appearances to mild flattening, reduction in size and fragmentation and this has a changeover range of age. For example, at the age of 2years 6 months, the capital femoral epiphyses may still be normal or be mildly abnormal. **Figure 4c (page 40)**
2. At about the age of 10, there is a change from severely flattened, fragmented, small capital femoral epiphyses to total absence. **Figures 4d,e (page 40)** The 'blurring' at the extremes of the age ranges of the expected features of a condition is a reflection of the normal variability found in evolving situations during growth and development.

Aspects of temporal reasoning in the diagnosis of skeletal dysplasias

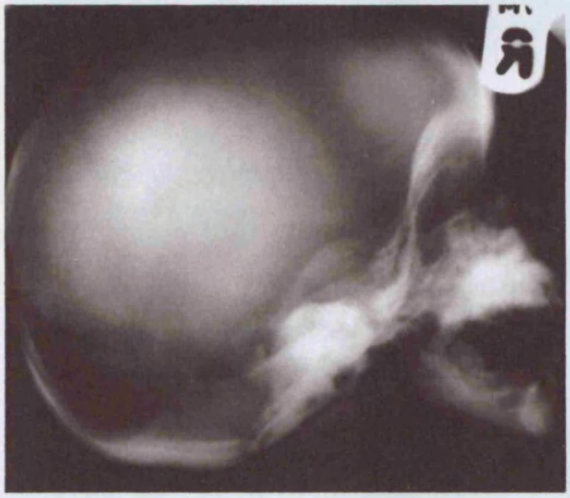
Table X



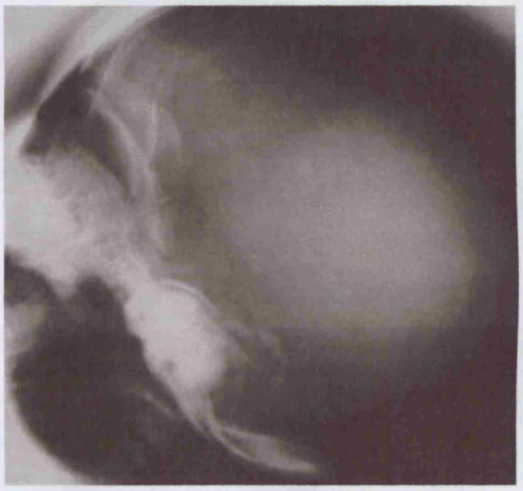
Craniometaphyseal dysplasia and craniodiaphyseal dysplasia show virtually indistinguishable appearances of the skull in infancy with sclerosis and thickening of the vault and base **Figures 5a,b** . In craniodiaphyseal dysplasia this progresses to striking sclerosis and overgrowth **Figures 5d,f**, but in craniometaphyseal dysplasia the changes gradually resolve **Figure 5c** and by later childhood the skull vault is of normal thickness and only minimal sclerosis along the suture lines **Figure 5e**. The ability to assess specific features of a condition changing with time implies knowledge of the natural history of the disorder. This information is not always available, especially for those rare dysplasias and malformation syndromes with few reported cases.

Figures 5a,b

Craniometaphyseal dysplasia (infant)

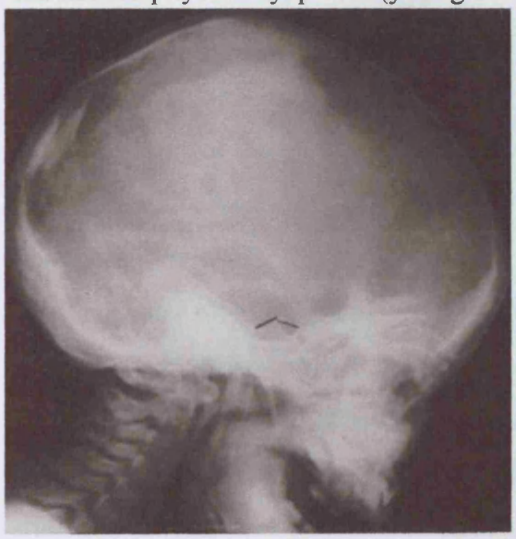


Craniodiaphyseal dysplasia (infant)

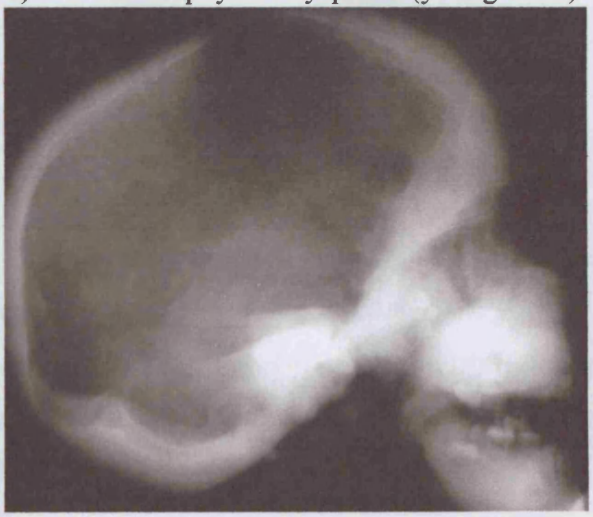


Figures 5c,d

Craniometaphyseal dysplasia (young child)

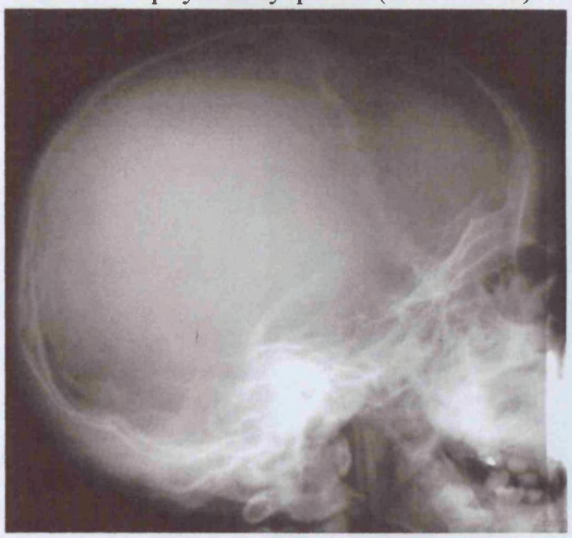


Craniodiaphyseal dysplasia (young child)

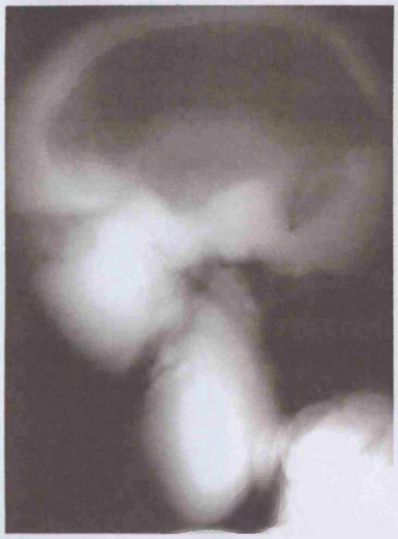


Figures 5e,f

Craniometaphyseal dysplasia (older child)



Craniodiaphyseal dysplasia (older child)



The following types of temporal information form an essential part of the diagnostic process:

a. Changes to normal processes

Skeletal dysplasias may cause normal processes of development to occur outside the expected age ranges, or to have a longer or shorter duration than normal. The effects may be general or localised. The findings of premature or delayed maturation, or increased or retarded growth velocity, may be significant in establishing a diagnosis.

i. Changes in maturation

For an individual dysplasia, a feature may be that of a generalised delay in bone maturation. Depending on the age of the patient, this statement would be taken to mean that epiphyses which should be ossified were not, or that if ossified, they were smaller than expected, or that epiphyseal plates which should have fused had not yet fused. For example in multiple epiphyseal dysplasia there is a generalised delay in bone maturation with epiphyses and carpal bones being smaller than expected for a given age. Alternatively there may be a localised abnormality of maturation, for example, the short-rib-polydactyly group of conditions show premature ossification of the capital femoral epiphyses as shown in asphyxiating thoracic dystrophy **Figure 6**, and upper

humeral epiphyses in short rib polydactyly syndrome type II, with ossification being present at birth (not present normally until several months of age). **Figure 7**

Figure 6 Advanced ossification

Asphyxiating thoracic dystrophy - neonate

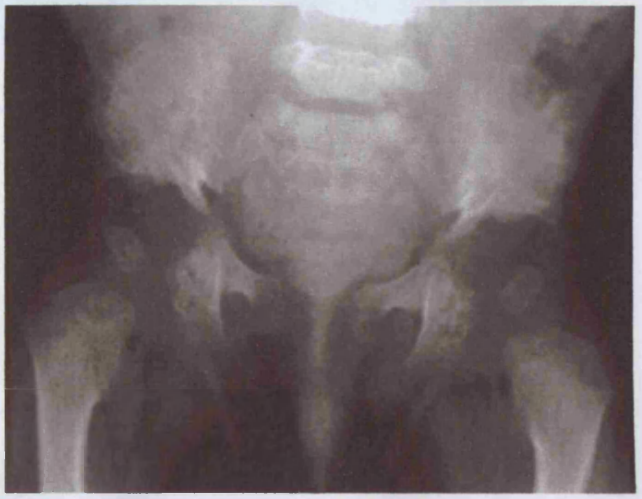
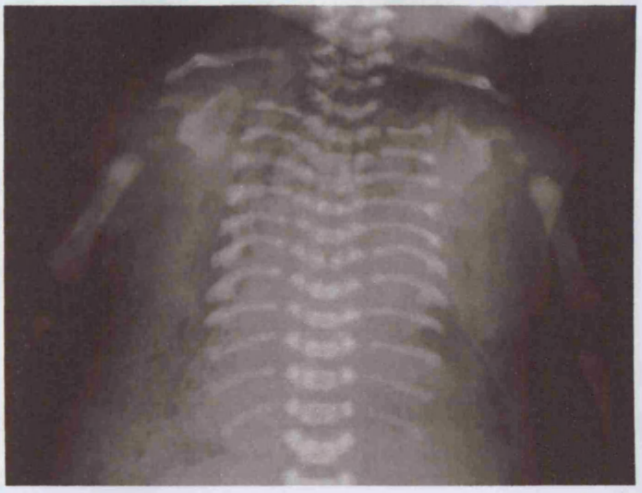


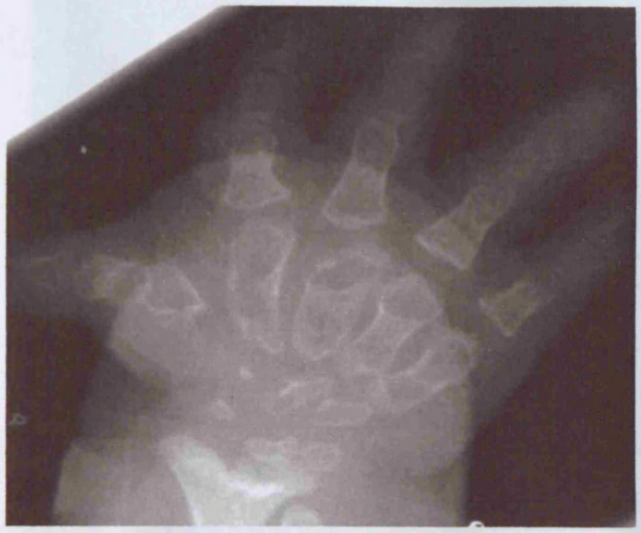
Figure 7 Short rib polydactyly syndrome type II - stillbirth



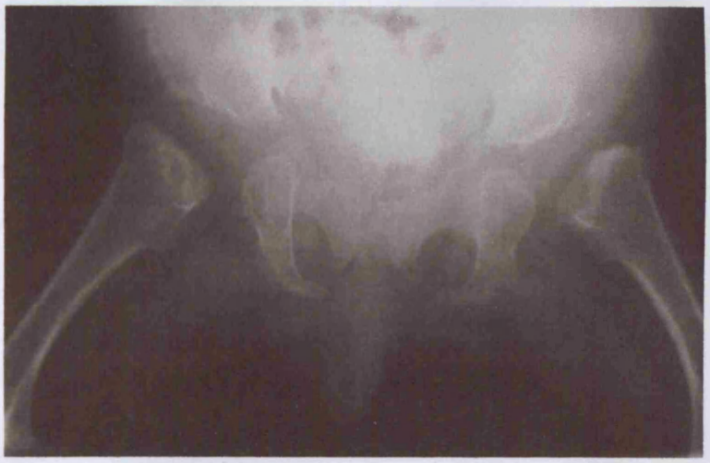
Dysharmonious maturation may be present, for example in diastrophic dysplasia, with selectively advanced ossification of the carpal centres (only two would be expected) but delayed ossification of the capital femoral epiphyses. **Figures 8a,b**

Figures 8a,b Diastrophic dysplasia

1 year



2 years



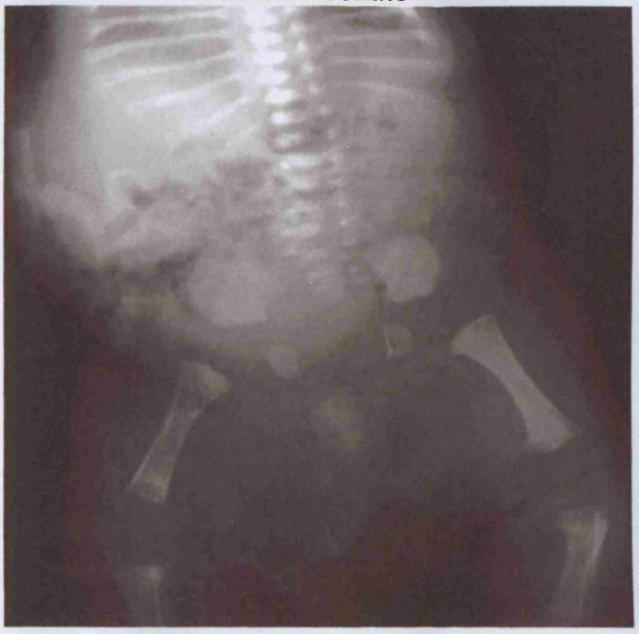
Absent (or delayed) ossification of the knee epiphyses and the pubic rami is a feature of neonatal spondyloepiphyseal dysplasia congenita (SEDC). Normally these would be ossified at birth. **Figure 9**

ii. Changes in growth velocity

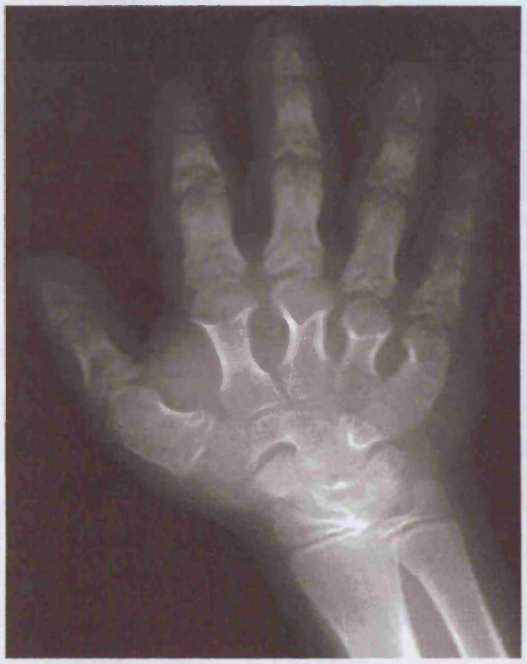
Localised abnormalities of skeletal maturation often result in localised changes in growth velocity. For example, premature fusion of epiphyseal growth plates is found in the presence of cone-shaped epiphyses in the hands and feet. This results in decreased or absent growth velocity in these areas such as is seen in pseudohypoparathyroidism, pseudoachondroplasia, acromesomelic dysplasia and acrodysostosis. These conditions have significant shortening of metacarpals and phalanges through childhood, but the premature fusion of cone-shaped epiphyses means there is subsequently significant deceleration of growth and progressively relatively more severe shortening. **Figure 10**

Figures 9 and 10

SEDC neonate

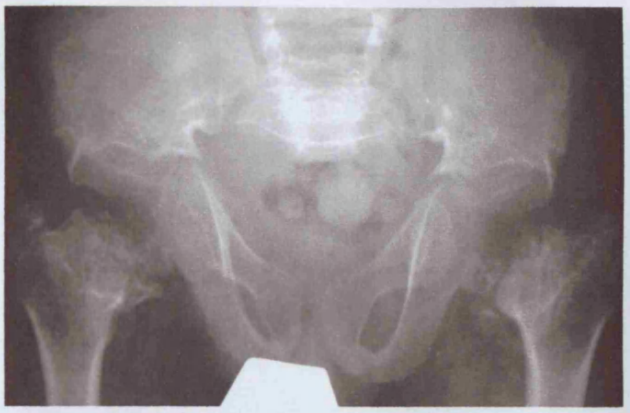


Acrodysostosis 9 years

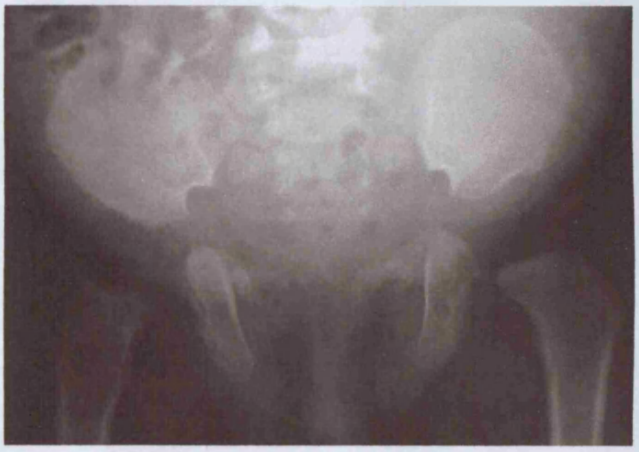


SEDC also shows changes in growth velocity. Two types of SEDC are recognised, one with severe coxa vara and one with mild coxa vara. The two types are indistinguishable at birth and up to about two years of age. Thereafter, there is a difference in growth velocity. SEDC without severe coxa vara runs parallel to, but below the third centile with a final predicted height of 140cm whereas SEDC with severe coxa vara shows a fall-off of growth velocity with a final height of 120cm. This final difference in height is over and above the localised shortening from the coxa vara [36]. **Figures 11a,b**

Figures 11a,b SEDC with severe coxa vara – 5 years



SEDC with mild coxa vara – 4 years

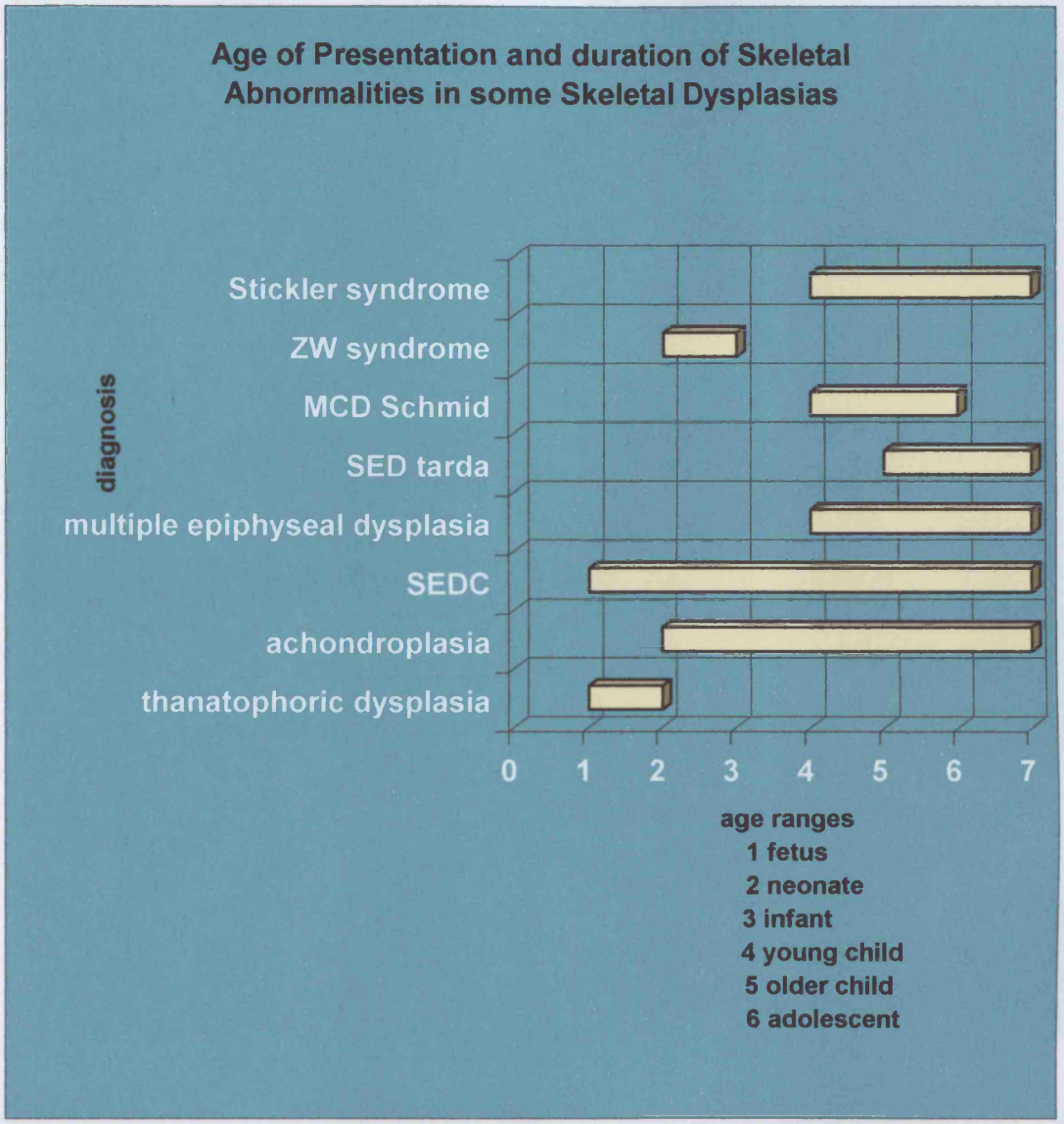


b. Onset and duration of the condition

The age of presentation of each condition varies **Table XI.** Some may be identified on the basis of specific malformations (such as short limbs, or a narrow thorax) as early as 18 weeks gestation on ultrasound scanning. Some dysplasias, such as multiple epiphyseal dysplasia (MED), present in early childhood, and others, including spondyloepiphyseal dysplasia tarda (SEDT), not until late childhood. In practice, this means that when considering the radiographs of a neonate presenting with a dysplasia or malformation syndrome, conditions presenting later can be excluded from consideration, so limiting the diagnostic possibilities. The reverse situation does not necessarily apply. When evaluating radiographic findings in childhood, the information is not always available as to the earliest age the findings were apparent and therefore earlier presenting conditions may still need to be considered. However, in this situation, perinatally lethal conditions can be excluded from consideration.

Conditions also vary in duration; for example, a short-rib-polydactyly syndrome presents at 18 weeks gestation and is stillborn, MED presents from 5 years with normal life expectancy; achondroplasia presents from 25 weeks gestation and has a normal life expectancy.

Table XI



c. Onset and duration of particular abnormalities

Different dysplasias may cause abnormal processes to start at different times, and have different durations.

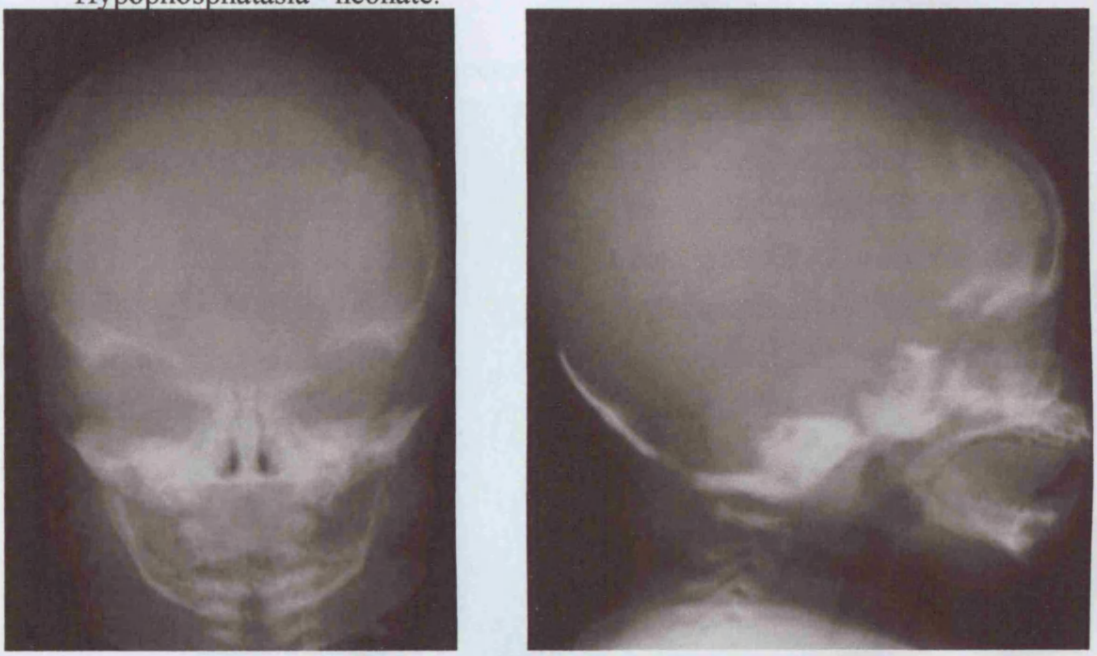
Those disorders with a short natural history (such as thanatophoric dysplasia, which is perinatally lethal, and SEDT, which only presents in late childhood) show only a minor progression of changes in the individual diagnostic features. For example, in thanatophoric dysplasia, the short ribs, short, bowed femora and platyspondyly are constant and invariable diagnostic findings.

However, in dysplasias with a longer natural history, extending prenatally (e.g. SEDC) or from birth (e.g. achondroplasia) to adulthood, with near normal life expectancy, the

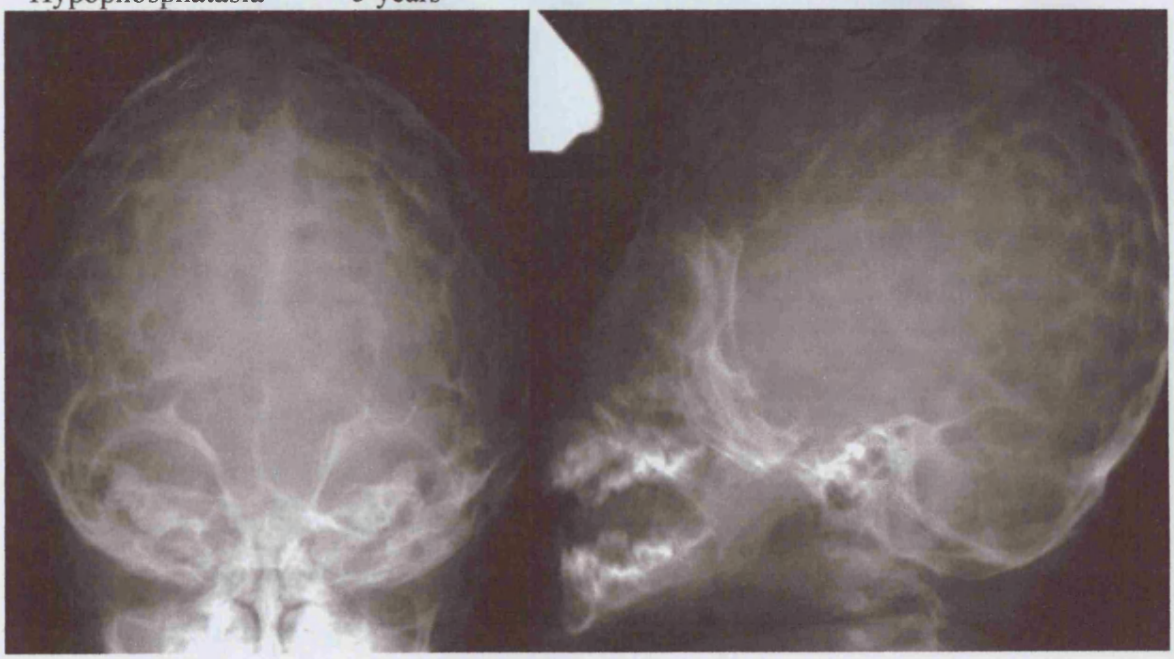
individual diagnostic features change with increasing age. Features appear, develop, and may disappear, either as a result of the process of maturation, or as part of the disease process. Features may also change in appearance and in degree of severity. For example the appearances of the skull in hypophosphatasia show a striking change over a short space of time, from the neonatal poorly ossified skull with wide sutures to craniostenosis due to premature fusion of the sutures and a copper-beaten appearance in early childhood shown in **Figures 12a,b,c,d**

Figures 12a,b,c,d

Hypophosphatasia - neonate.



Hypophosphatasia 5 years

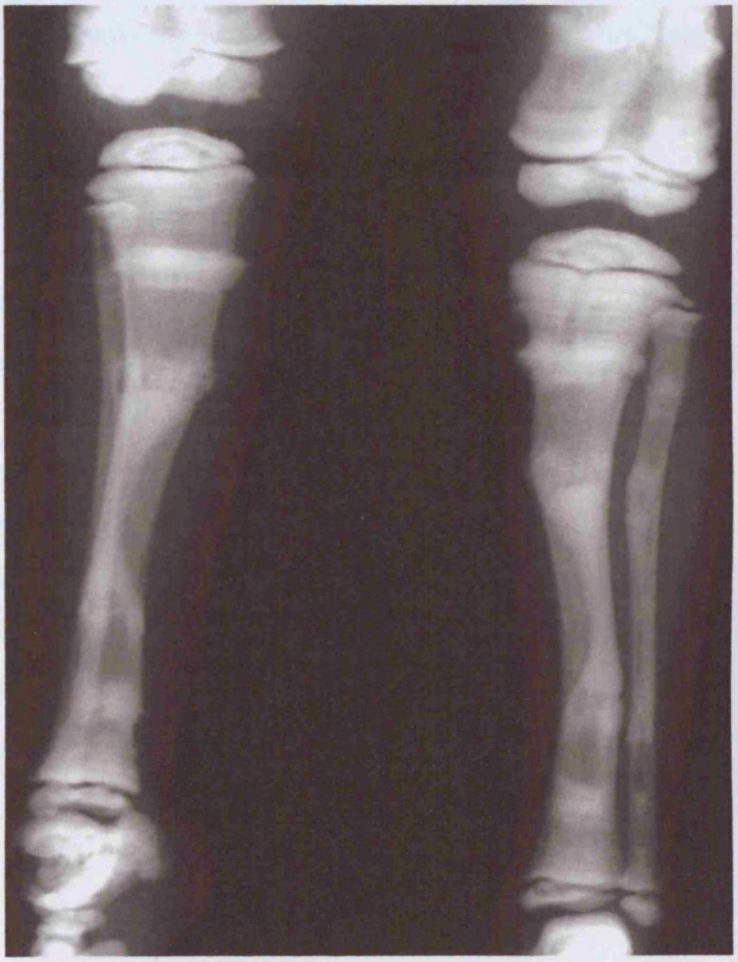


d. Evidence of previous abnormal processes

Some abnormal processes leave evidence that is visible even once the bone is fully matured. For example, the natural history of some disorders involves changes of bone-density with time. This is exemplified in the juvenile (intermediate) form of osteopetrosis. In the neonate there is dense sclerosis (sometimes also seen in normal neonates). Subsequently there are periods of normal bone growth followed by abnormal bone growth resulting in sclerotic bands and abnormal modelling. The neonatal sclerotic blueprint of the bone remains visible throughout growth.

Figure 13

Osteopetrosis 12 years



Other abnormal features may disappear without trace, either as part of the abnormal process or when the normal skeletal maturation masks previously visible abnormalities. An example of the former is chondrodysplasia punctata in which the neonatal period is characterised by dense stippled calcification in the region of developing epiphyses. This stippling gradually disappears, usually within the first year. An example of the latter is

given by the metaphyseal chondrodysplasias. After closure of the epiphyseal plates in adulthood, the diagnostic changes at the metaphyses are no longer apparent.

Rarely, all the radiographic features of a condition may be absent (normal) for a period of time. For example, some patients with Stickler syndrome present in the neonatal period with rhizomelic limb shortening and broad metaphyses of the long bones (dumbell-shaped). There is an associated Pierre Robin anomaly. This, neonatal expression is known as the Zweymuller-Weissenbacher (ZW) syndrome. By the age of one year, there are no abnormal radiographic findings. During mid-childhood (5-6 years) typical features of Stickler syndrome develop with small, flattened irregularly ossified capital femoral epiphyses and localised platyspondyly. This interposed period of normality is an uncommon manifestation of dysplasias and malformation syndromes in general.

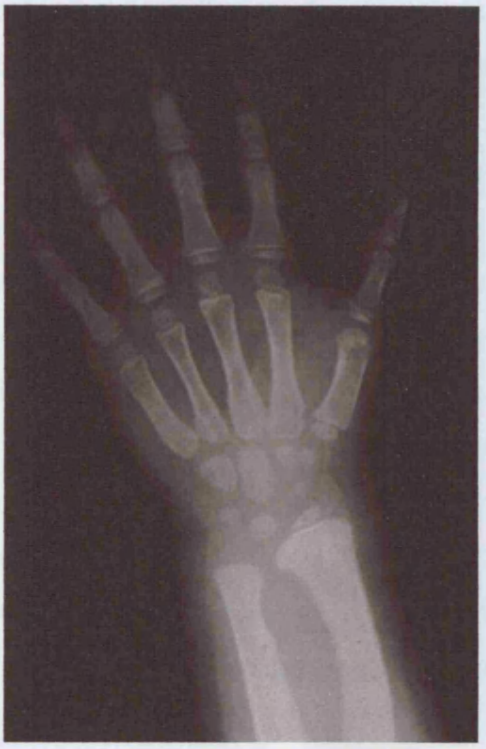
e. Relationship between abnormal processes and final deformities

Causal relationships between underlying processes and resultant deformities may mean that expected features of a dysplasia are not apparent until the underlying processes are complete.
For example, dyschondrosteosis and Turner syndrome may have a Madelung deformity. This results from premature fusion of the medial side of the distal radial epiphyseal plate, with subsequent growth of the lateral side of the radius and of the ulna, resulting in a decrease in the carpal angle and dislocation of the distal end of the ulna. Madelung deformity is not seen until later childhood, when this localised physeal fusion occurs.
Figures 14a,b,c,d

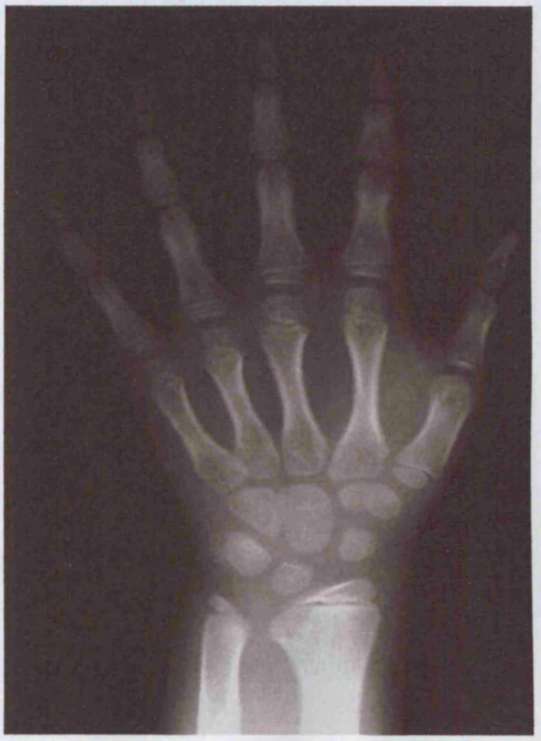
Figures 14a,b,c,d

Dyschondrosteosis

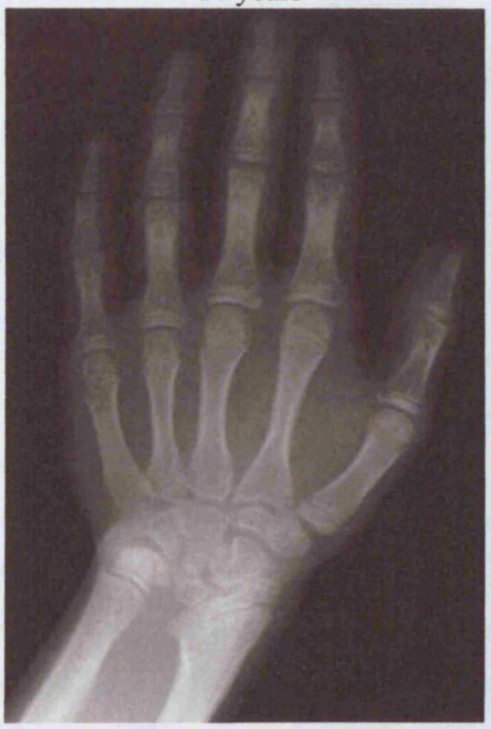
7 years



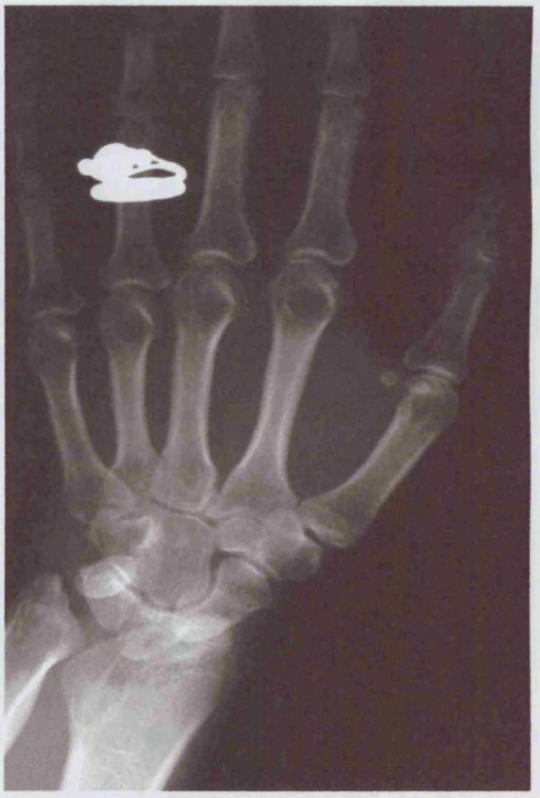
9 years



14 years



adult



Chapter 4. The Expert System in Practice

The clinical user enters simple clinical information and radiological findings from the skeletal survey of the case under consideration into the system at the initial consultation stage. More information can be offered at later stages of the consultation. In addition the system asks questions to elicit further information. Each case finding is categorised by the system as either a **hard** or **soft** finding. **Hard** findings are significant abnormalities which must be adequately accounted for by an acceptable diagnosis. **Soft** findings may be attributable to neutral or natural causes, such as 'broad thorax' or 'prominent eyes', or they may be mild findings that are difficult to differentiate from normal findings. They are weighted differently in any scoring system.

Conditions generated for consideration are those in which a **primary trigger** (all are hard findings) is matched. There will be several possibilities at this stage (a list of gamuts). The initial selection of diagnoses for consideration is determined by ranking the possibilities according to how well they account for the common (expected) case findings. The system explores the initial set of possibilities by asking questions about their unknown **sufficient set** and **typical** features. The user is asked to answer 'yes', 'no', or 'unknown'.

After this initial exploration the possibilities are re-evaluated and the possibility that satisfies a reasonable subset, the current set of hard abnormalities, becomes a working hypothesis. Conditions that no longer satisfy the set of hard abnormalities are demoted. The current differential is recomputed after every diagnostic cycle and therefore a hypothesis may enter, leave and re-enter the focus. Normally there are two or three working hypotheses which are evaluated very closely. This is done by looking at entered findings which have not been used, checking the hypothesis' expectations not yet observed (re-examining available radiographs), finding currently unavailable information (obtaining more radiographic views), or confirming the accuracy of particular findings.

The differentiation strategy uses **secondary triggers** or differential diagnostic features to differentiate between a small group of highly promising possible diagnoses. Secondary triggers are also brought into play when the exploration of a working hypothesis identifies unexpected findings, not in keeping with the hypothesis under consideration. The unexpected finding may suggest another possible diagnosis through the secondary triggers. Dysplasias associated through secondary triggers tend to share many of their features, which means that they are likely to be confused with each other.

A hypothesis may be concluded if it is good enough in absolute terms, which includes adequate coverage of hard abnormalities, and if it is sufficiently better than the next-best competitor.

If a hypothesis is exhaustively explored and no firm decision can be reached, then the hypothesis becomes explored and is ranked in order of probability with other hypotheses. If a diagnosis cannot be concluded as a result of insufficient information that subsequently becomes available, the hypothesis may be reinstated as a possibility.

Please see Appendix B for a sample consultation with the expert system.

Chapter 5. Evaluation of the Prototype Expert System

The initial trial aimed to assess the hypothesis that general radiologists achieved poor diagnostic accuracy in the field of skeletal dysplasias and malformation syndromes using standard methods of diagnosis, and also that an expert system could improve their accuracy [27, 28, 29].

Method

From the dysplasia frames incorporated in the system, a sub-set of 35 was identified, all with 'platyspondyly' as a feature. This common feature was a **primary trigger** and therefore all 35 conditions would be brought forward for diagnostic consideration at the initial consultation. This sub-set was used because all related conditions, including the frames for differential diagnoses had been incorporated. 10 skeletal surveys were selected from within this sub-set for the trial and two experts in the field verified the diagnoses. Six general radiologists took part in the trial – none was an expert. Three were asked to arrive at a diagnosis using standard methods of textbooks and journals and three were asked to arrive at a diagnosis using the expert system. Unlike normal radiological practice, no time constraints were imposed. Overall there were 18 attempts at diagnosis using textbooks and 34 attempts using the expert system.

Results

Of the 18 textbook diagnoses only one was correct. In some cases there was a misdiagnosis in spite of all the relevant case-findings being observed. **1/18**
A diagnosis was reached using text books (although a misdiagnosis) much faster than by using the expert system.
21 out of the 34 diagnoses using the system were correct. The system in fact reached the correct diagnosis from the findings identified during the textbook evaluation, which had resulted in misdiagnosis. **21/34**

Discussion

A detailed evaluation of the individual case findings identified problem areas.
Some cases misdiagnosed by the system were attributed to user misinformation. This is a problem for any consultation system that relies on human input.

Some participants using the system were reluctant to give specific information in case it proved to be wrong, falling back instead on more general descriptions such as 'the spine is abnormal' and answering 'unknown' to more detailed questioning. A major problem in diagnosis is recognising unfamiliar radiological patterns and appearances. One reason why textbooks are of little assistance is that only a limited number of illustrations can be published. There is evidence that in radiology, the use of images to guide and validate user input can improve the accuracy of information given. It was recognised that the incorporation of images could help the user to identify features which were either rare or subtle and could be confused with other features or with normality.
In other cases where the system failed to make an accurate diagnosis, it homed onto a particular diagnosis and asked leading questions early in the consultation when there was not sufficient evidence to do this.

Conclusion

The preliminary results confirmed the impression that general radiologists achieve poor diagnostic accuracy in the field of skeletal dysplasias when using standard methods of diagnosis. The use of the prototype expert system dramatically improved diagnostic accuracy in this field.
As a result of these trials refinements to the diagnostic system were made. These included changes to the diagnostic engine to provide a more sophisticated system of evaluation with more appropriately weighted features and refinement of the scoring system.

Chapter 6. Clinical Trial: comparison of the Expert System with Standard Methods of diagnosing Skeletal Dysplasias by General Radiologists.

Aim

The purpose of the clinical trial was to test the hypothesis that the computerised knowledge-based system could achieve an improvement in diagnostic accuracy by general radiologists in the field of skeletal dysplasias and malformation syndromes in children. This was following implementation of refinements to the prototype system and expansion of the knowledge base with the incorporation of more dysplasia frames.

Background

A retrospective analysis of all skeletal surveys referred for a second opinion to the department of radiology at Great Ormond Street Children's Hospital, London, during a one-year period was conducted. This is described in detail earlier. The study included an estimate of accuracy of the referral diagnosis, broken down by specialty. The diagnostic accuracy of all the referrals was 46%. An estimate of the diagnostic accuracy of radiologists was 31% compared with an estimated accuracy of 57% by clinical geneticists.

6.1 Design, Material and Methods

The expert system is interactive, prompting the user (radiologist) to enter a short initial list of major findings identified from a radiographic skeletal survey, and then using these to trigger possible diagnoses. These are further explored by asking the user targeted questions relating to the radiographs. The user may enter additional findings at any stage, and the system brings new possibilities into consideration in the light of new information (both responses to queries and user-initiated entries). The trial was not designed to test the 'user friendliness' of the system and no time constraints were imposed.

The design of the trial was discussed and refined on the advice of Professor A. P. David, Professor of Statistics, UCL.

A sub-group of the sclerosing and cranio-tubular dysplasias was selected from the whole group of skeletal dysplasias. The dysplasia frames relating to this sub-group had been completed and included frames for any differential diagnoses which may have been needed to be considered.

Eighteen different conditions were included. These included relatively common conditions such as osteopetrosis and osteopoikilosis and much more rare conditions such as Pyle's disease (metaphyseal dysplasia), craniodiaphyseal dysplasia, oto-palato-digital syndrome type II. Thirty-two skeletal surveys were chosen, and their diagnoses verified by a panel of three or four radiologists experienced in the field of skeletal dysplasias. The panel consisted of Dr Donald Shaw (London), Professor Andres Giedion (Zurich), and Professor Alan Oistreich (Cincinnati) and myself. The skeletal surveys were evaluated independently. Consensus agreement could not be reached because of time constraints with my foreign visitors. However Dr Shaw and I had agreement in all cases and all were considered typical examples.

100% agreement in **21/32**
Majority agreement in **6/32**
50% agreement in **5/32**

(One case, hypophosphataemic rickets, had been substituted for another on the basis of insufficient agreement and uncharacteristic findings.)

Eight groups of skeletal surveys, with eight cases in each group, were created for examination by eight radiologists. Each radiologist examined the skeletal surveys of two groups, sixteen cases altogether, one group of eight skeletal surveys using standard interpretation with the help of reference textbooks and the other eight using the system. A wide range of specialist, up to date textbooks and appropriate journals were available for consultation and no time restraints were imposed. Each case was thus evaluated four times by four different radiologists, twice with textbooks and twice using the system. Each group did not include more than one case with a particular diagnosis, and no individual radiologist saw the same case more than once. However one diagnosis may have occurred twice within the two groups examined by each radiologist.

To evaluate the possibility of a learning process, half of the radiologists were asked to examine the skeletal surveys using textbooks first and half using the diagnostic system first. In addition there was a time interval of at least one week between reporting on the two different groups.

Table XII shows the 8 Groups (labelled A-H) and the number assigned to each skeletal survey.

Group	Case numbers						
A	1	5	9	13	17	21	26	29
B	2	6	10	14	18	22	25	30
C	3	7	11	15	20	23	27	31
D	4	8	12	16	19	24	28	32
E	3	6	11	14	20	22	25	30
F	4	8	9	16	17	24	26	29
G	1	5	12	13	19	21	28	32
H	2	7	10	15	18	23	27	31

All the radiologists for this study were those who had recently completed their final FRCR examination. It was considered that they would have read the most recent literature and would have had recent teaching on the subject and studied some texts on skeletal dysplasias in their examination preparations. They would be most likely to arrive at a correct diagnosis.

With each skeletal survey there was some basic clinical information which included patient age, age at presentation and gender. Other clinical findings included deafness, partial alopecia, large jaw, abnormal facies or cleft palate, but the majority did not have further clinical information. This was considered to be the information generally available from the request form generated for a radiographic skeletal survey.

Each radiologist was asked to identify between five and ten major abnormalities at the beginning of the consultation with the system.

Table XIII

Groups of 8 skeletal surveys (A-H) examined by 8 radiologists

Radiologist	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Standard	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**
textbook	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Diagnosis								
Diagnostic	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**
system	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**A**

6.2 Results

The system presented its conclusions as a list of ranked possible diagnoses. If one diagnosis scored both sufficiently highly and was significantly higher than any competing diagnosis, (using predetermined threshold scores) the system would conclude this as the correct diagnosis. If no diagnoses reached a pre-determined threshold for consideration as a possibility, the system would conclude that it could not make any suggestions.

When using the books, the radiologists were asked to list the major findings, followed by either their conclusion, or a list of suggested diagnoses in order of likelihood, or, if they could find no suitable candidates, to state 'unknown'.

The results were classified as follows:

Correct top

System either concludes the diagnosis and ends the consultation, or ranks the diagnosis first out of a list of differential diagnoses.

Books the correct diagnosis is firmly established or ranked top of a list of differential diagnoses.

Correct suggested

System the correct diagnosis is not ranked first, but is in the top three suggested diagnoses.

Books correct diagnosis listed in differential diagnosis, but not top.

Unknown

System no diagnoses scored sufficiently highly to be suggested as possibilities.

Books no diagnoses suggested; radiologist states 'unknown'.

Incorrect suggested

System the correct diagnosis does not appear in the top three possibilities.
Books the correct diagnosis is not suggested.

Incorrect concluded

System an incorrect diagnosis is concluded.
Books an incorrect diagnosis is concluded.

Table XIV - Table of Results

case	dysplasia	correct top		correct suggested		unknown		incorrect suggested		incorrect concluded	
		ES	book	ES	book	ES	book	ES	book	ES	book
1	Pycnodysostosis	2	1				1				
2	Osteopetrosis	2	2								
3	Craniodiaphyseal	2	2								
4	Craniometaphyseal	2	1						1		
5	Epiphysealis Hemimelica	2	2								
6	Progressive Diaphyseal	1	2	1							
7	Frontometaphyseal	1	1					1			1
8	Melnick-Needles	2	1								1
9	Osteopathia Striata Cranial involvement	2	1								1
10	Craniodiaphyseal	2	1								1
11	Osteopetrosis	1	2					1			
12	Progressive Diaphyseal	2	2								
13	Craniometaphyseal		1	1						1	1
14	Osteopoikilosis	2	2								
15	Pyles			2							2
16	Pycnodysostosis	2	2								
17	Osteopetrosis	1	1	1							1
18	Endosteal Hyperostosis	1	1					1	1		
19	Osteopetrosis	1	1					1			1
20	Osteopathia Striata Cranial involvement	2							1		1
21	Sclerosteosis	1	1							1	1
22	Pyles	2			1						1
23	Melorheostosis			1				1	1		1
24	Frontometaphyseal	1	1		1					1	
25	Epidermal Naevus	1		1					2		
26	OPD II	2									2
27	OPD I	1		1					1		1
28	Craniometaphyseal	2							2		
29	Tuberous Sclerosis	1	2					1			
30	Pycnodysostosis	2	1						1		
31	Epidermal Naevus	1		1					1		1
32	Osteopathia Striata Cranial involvement	2	1						1		
Total		46	32	9	2	0	1	6	12	3	17
%		72%	50%	14%	3%	0	2%	9%	19%	5%	27%

Correct identified or suggested by the system = 55/64 86%

Correct identified or suggested by books = 34/64 53%

Incorrect identified or suggested by the system = 9/64 14%

Incorrect identified or suggested by books = 29/64 45%

Table XV.

Results by radiologist

Radiologist	correct system	incorrect system	correct books	incorrect books
1	8		6	2
2	6	2	6	2
3	8		3	5
4	8		5	3
5	6	2	5	3
6	6	2	4	4
7	6	2	4	4
8	7	1	1	7

Radiologist 2 showed no change in diagnostic accuracy but all other radiologists reached more, accurate diagnoses when using the system.

The trial results of the 5 cases in which there was only 50% agreement by the experts was reviewed. These were case numbers 16, 18, 22, 24 and 25.

Correct identified or suggested by the system = 8/10 (80%)

Correct identified or suggested by books = 6/10 (60%)

Incorrect identified or suggested by the system = 2/10 (20%)

Incorrect identified or suggested by books = 4/10 (40%)

The results for these 5 cases are very much in line with the overall results of the trial and do not appear to have resulted in any form of bias

6.3 Analysis and Discussion

There was no significant difference between the results of radiologists using the books first and those using them second; similarly, there was no significant difference between those using the expert system first and those using it second. We can therefore conclude that there is no evidence of 'cross-over' learning.

The results can be divided into two categories of outcome:
Favourable , where the correct diagnosis is either concluded or in the top three
Unfavourable where an incorrect diagnosis is either suggested or concluded

As far as the patient is concerned, a favourable outcome enables an accurate diagnosis to be made, sometimes following further targeted specific diagnostic tests such as clinical and dysmorphic evaluation or biochemical or molecular studies. A firm diagnosis means that meaningful genetic counselling for the patient and their family can be offered. With an understanding of the natural history of the condition, complications can be predicted and sometimes prevented and management planned. In some conditions treatment may even be curative.
An error in diagnosis is an extremely unfavourable outcome for the patient and leads to inaccurate labelling and inappropriate management.

In terms of producing a favourable outcome, seven out of the eight radiologists produced more favourable outcomes using the system than using textbooks. The eighth (radiologist number 2) produced equal numbers of favourable outcomes using the system and using the books. One radiologist (radiologist number 8) showed a dramatic improvement in diagnostic ability, improving from seven out of eight inaccurate diagnoses using the textbooks to seven out of eight accurate diagnoses using the system. Overall, the system produced 55/64 favourable outcomes; using the books resulted in 34/64 favourable outcomes. We can therefore conclude that using the system significantly increased the likelihood of a general radiologist reaching a favourable outcome.

In terms of avoiding errors, seven out of the eight radiologists produced more errors when using the books than when using the system. Overall, the system produced 9/64 errors as opposed to 29/64 produced when using textbooks.

Statistical analysis of the results by a paired t-test showed them to be significant at $p<0.005$. The mean difference with 99% confidence interval is 2.5 +/- 1.89 (0.61-4.39). We can therefore conclude that the system is significantly more accurate.

The results of the individual cases were analysed to attempt to determine the reasons for failure to establish a diagnosis using the system and using books.
Three reasons for an inaccurate diagnosis were identified. Firstly, the inaccurate identification of findings; secondly, failure to identify the correct diagnosis as a possibility and thirdly failure to reach the correct conclusion despite considering the correct diagnosis as a possibility. These three occurred whether using the system or books.

Analysis of the cases showed that the percentage of inaccurate findings entered is not a good predictor of success or failure. However performance is affected by the misrecognition of certain highly significant findings (hard features). For example, in one attempted diagnosis of case 2, 50% of the observations were incorrect and 58% of the questions answered affirmatively were incorrect. In spite of this the system scored the correct diagnosis (pycnodysostosis) highest out of the six generated hypotheses because the incorrect findings were not highly significant (soft features) (e.g. some ribs hypoplastic; abnormal vertebral trabeculae; irregular acetabula). In one attempted diagnosis of case 11, all six observations were correct, as were three out of the four questions answered affirmatively but the correct diagnosis was not triggered by the system because the user had failed to recognise any sclerosis, a significant finding. In the book diagnoses, for example where an incorrect conclusion was reached, the percentage of accurate observations varied between 40% and 100% (average 80%), but it was still possible to reach a correct diagnosis with only 40% accurate observations.

The system does not evaluate radiographs directly and accepts the radiologist's assessment of the images. This is also true with textbook evaluation. Future development will link the expert system to an image database of radiographs, which will provide images for the radiologist to use as a comparison. This should help to reduce inaccurate identification of findings. It might also be possible to identify groups of findings, which are particularly difficult to report accurately and use images matched to those findings to help the user.

There was a significant difference between the results using the textbooks and those using the system with regard to the generation of hypotheses. The system only failed to generate the correct hypothesis in three cases (5%), whereas radiologists using the books failed in thirty cases (47%). Of these cases, where the findings observed by the radiologists were then entered onto the system, the system successfully generated the correct hypothesis in twenty-seven of the thirty cases. This may partly represent a difference in approach; the radiologists in general presented their results as a conclusion, only rarely listing possible alternatives, whereas the system is designed to generate multiple hypotheses and then explore and rank them. This has the advantage of suggesting hypotheses, which may otherwise not have been considered and may be particularly helpful in the case of more rare dysplasias, such as Pyle disease, OPD-I and OPD-II and epidermal naevus syndrome, none of which were diagnosed by the radiologists using textbooks. The system suggested the correct diagnosis in all four attempts at Pyle disease, in both attempts at OPD-II, both attempts at OPD-I and in all four attempts at epidermal naevus syndrome.

The background audit of cases referred for a second opinion identified an accuracy of 31% by referring radiologists. In this clinical trial the accuracy was 53% using standard methods of diagnosis. The difference may be related to the fact that fewer straightforward (common) conditions are referred. In fact the more rare diagnoses are more likely to be misdiagnosed overall.

A rough and ready combined assessment of prevalence of the individual conditions used in th trial has been made. This has been based on my patient database and personal experience and also that of Lachman (as reported in Taybi and Lachman up to 1994) [11]. Taybi gives an additional evaluation based on literature reports. These are slightly misleading in some instances as more florid conditions with striking clinical changes are more frequently reportec (for example epidermal naevus syndrome). The comments in brackets are my evaluation of th four conditions in which Lachman does not give personal experience figures. These may be a variance with the literature reports but this is because, in these malformation syndromes, skeletal changes are a variable (uncommon) manifestation.

Table XVI

An estimate of prevalence of the individual conditions used in the trial.

Diagnoses used in the Trial	Patients on my database	Literature	Patients seen by Lachman	combined ranked evaluation
osteopetrosis	46	750	15	61
osteopoikilosis	12	common		(common)
tuberous sclerosis	6	common		(common)
craniometaphyseal dysplasia	10	85	7	17
Melnick-Needles syndrome	7	45	8	15
dysplasia epiphysealis hemimelica	9	160	5	14
diaphyseal dysplasia	7	170	7	14
pycnodysostosis	8	150	6	14
Pyle's disease	6	35	4	10
OPD type II	1	20	7	8
osteopathia striata	7	35	1	8
melorheostosis	2	320	5	7
craniodiaphyseal dysplasia	5	25	1	6
frontometaphyseal	4	35	2	6
epidermal naevus syndrome	5	230		(rare)
OPD type I	2	100		(rare)
sclerosteosis	1	65	0	1
endosteal hyperostosis	1	50	0	1

Rough assessment of the prevalence of dysplasias from the combined personal experience of myself and Lachman.

Prevalence	**Number of cases seen**
Common	14-70
Intermediate	6-13
Rare	0-5

Table XVII

Comparison of diagnostic accuracy between commonly encountered dysplasias, those of intermediate frequency and rare conditions.

	Expert system correct top or suggested	Books correct top or suggested
Common		
osteopetrosis	6	6
craniometaphyseal dysplasia	5	2
Melnick-Needles syndrome	2	1
dysplasia epiphysealis hemimelica	2	2
pycnodysostosis	6	4
diaphyseal dysplasia	4	4
tuberous sclerosis	1	2
osteopoikilosis	2	2
	28	**23**
Intermediate		
OPD type II	2	0
osteopathia striata/cranial involvement	6	2
melorheostosis	1	0
frontometaphyseal	2	2
craniodiaphyseal dysplasia	4	3
Pyle's disease	4	1
	19	**8**
Rare		
OPD type I	2	0
epidermal naevus syndrome	4	0
endosteal hyperostosis	1	1
sclerosteosis	1	1
	8	**2**

These findings confirm that the expert system significantly enhances diagnostic accuracy in conditions which are less frequently encountered (the 'intermediate' and 'rare' groups) when compared with the minor improvement in diagnostic accuracy in the 'common'

group. These findings explain the relative improvement in diagnostic accuracy using textbooks during the trial (53%) compared with the accuracy achieved in the background evaluation of referred cases for a second opinion (31%).

When using books, the radiologists gave a ranked list of possibilities or a 'differential diagnosis' in only fourteen of the sixty-four diagnoses attempted (22%). In the other cases they either stated 'unknown' or gave a single diagnosis. Of the fifty single diagnoses given, twenty-seven were correct (54%). The system reached a specific conclusion in twenty-six cases (41%). Of these twenty-two were correct (85%). At present the system is weighted against reaching an absolute conclusion, in that it will only conclude a diagnosis when an hypothesis scores sufficiently highly in absolute terms and also has a score which is sufficiently higher than its closest competitors. The only exception to this is where a **'typical'** match is found. Most dysplasias do not have **typical features** and only two of the diagnoses reached by the system were based (correctly) on the match of a **typical feature**.

A clinical trial of OSSUM [24] reported that the correct diagnosis was identified an equal number of times when using the standard textbook method of diagnosis or when using OSSUM. OSSUM is a specialised database of radiological images for the diagnosis of skeletal dysplasias [17]. It is not an expert system but relies on cross-referencing all reported syndromes using a standardised list of clinical and radiological findings. The trial protocol was slightly different (using experienced paediatric radiologists) and the sample size small, but the authors observed an accuracy of 68% using either method. They further suggested that in practice a diagnosis was achieved following consultation rather than by an individual paediatric radiologist, and that both textbooks and databases/expert systems would be used. By combining the results from the use of textbooks and OSSUM the correct diagnosis was suggested (in the top three differential diagnoses) in 90% of cases, although the authors could not predict that the consultation process would result in the correct diagnosis being identified. The trial of OSSUM was conducted in a large department of paediatric radiology in the USA where the consultation process may be more feasible than in the UK, where there are only a handful of centres with three or more paediatric radiologists.

If similar 'consultation' criteria are applied to the trial of the expert system, then the correct (top) diagnosis using the system was identified in 30/32 cases (94%), The

correct diagnosis was in the top three suggested diagnoses in 32/32 cases (100%). When using textbooks the correct diagnosis was made in 24/32 cases (75%) and was included in the top three suggested diagnoses in 25/32 cases (78%).

The combination of the correct (top) diagnoses using both the expert system and textbook diagnosis did not further enhance the diagnostic accuracy above 94%.

6.4 Conclusions

As a result of these trials further refinements to the expert system will be implemented.

- Clearer instructions on the type and amount of information entered during the first phase of the consultation, when the radiologist is entering the initial findings.
- Modifications to the diagnostic engine to allow the system to explore more fully closely scoring competing hypotheses before reaching a conclusion.
- Link the system to an electronic image database to enable users to view age-matched images of a particular finding, which will help in recognition of abnormalities which are rare, complex, or easily misinterpreted and should help to reduce the problem of inaccurate input. It will also help by providing an illustrative image, when the terminology used has been unfamiliar, forcing the user to answer 'unknown' to a specific query.

However, even without these modifications the system has demonstrated its ability significantly to improve the performance of a general radiologist in the diagnosis of skeletal dysplasias and malformation syndromes.

In clinical practice, to further improve diagnostic accuracy the highest ranked diagnoses should be considered for subtle differentiating clinical and dysmorphic features and targeted biochemical and molecular studies.

Chapter 7. The development of a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias REAMS

To provide reference images to be linked with the expert system, a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias (REAMS) [15] was created. This database was developed to be a stand-alone product (without the expert system) and was created as part of the Oxford Medical Database series (including the London Dysmorphology Database – LDDB [19]), with the support of Oxford University Press. It was designed to have the same look and feel as other databases in the series using a windows format and software developed by Oxford University Press. It was recognised that although the aim was to link with the expert system, the database could be used as an aid to diagnosis in its own right. Additional software was created (by John Washbrook) to provide powerful search facilities using different parameters and to improve user friendliness and broaden or narrow searches and to create different reporting formats.

As other databases, it may be used as a reference source for teaching and learning in addition to providing an aid to diagnosis. When being used to help in diagnosis the limitations need to be recognised. It cannot cover all known conditions. Rare conditions have only a few images and therefore are incompletely covered. The full range of features expected in a condition may not be illustrated. In a few cases a finding (an abnormality identified on an image) may be different from an expected feature of an individual condition. In addition the database does not aim to provide a full skeletal survey for every case, only images showing the diagnostic features of the condition from a number of cases. Images of a condition will only cover the ages at which radiological features are present. The database requires the users to make their own observations and to undertake searches for potential diagnoses using their own decision-making. The user makes the final diagnosis from the suggested matches identified by the database. The user needs to have some experience both in the interpretation of radiographs and in the field of skeletal dysplasias.

I was responsible for selecting the images to be included on the database. They were chosen from my patient database of about 4000 cases. From these patients I reviewed 25000 images and selected 18000 for digitisation. The hardware (Vidar Systems Corporation VXR-12 film digitiser and computers) was provided by Oxford University

```
 B28045531              Last updated: 11-08-05 Created: 11-08-05 Revision: 1
LANG: eng           CAT DATE:    -  -    MAT TYPE: y         COUNTRY: enk
SKIP: 4             BIB LVL: m          BCODE3: -           MARCTYPE:
LOCATION: ull
008           s2005    enk||||f ma   00| 0|eng|dntm a
040     UkLU|beng|cUkLU
100  1  Hall, Christine Margaret.
245  14 The development and evaluation of two computer-based diagnostic aids
        in the field of inherited skeletal dysplasias and malformation
        syndromes.
300     122 leaves :|bill., charts.
500     CD-ROM in pocket at front.
500      Offprints of 4 published papers in pocket at back.
500     Leaves 99-122 are appendices.
502     Thesis(MD)--University of London, 2005.
691   1 Medicine (Subject Panel)
907     ag
981     |bT
```

```
 I30566927              Last updated: 11-08-05 Created: 11-08-05 Revision: 1
COPY #: 0           PATRON#: 0          RECAL DATE:    -  - INTL USE : 0
ICODE1: 0           LPATRON: 0          TOT CHKOUT: 0       COPY USE: 0
ICODE2: -           LCHKIN:    -  -     TOT RENEW: 0        IMESSAGE:
I TYPE: 5           # RENEWALS: 0       LOCATION: usthe     OPACMSG:
PRICE: #0.00        # OVERDUE: 0        LOANRULE: 0         YTDCIRC: 0
OUT DATE:    -  -   ODUE DATE:    -  -  STATUS: o           LYCIRC: 0
DUE DATE:    -  -   IUSE3: 0
098     MD 2005 UCL
BARCODE       1916960161
```

Press. The images were digitised at full resolution and archived in duplicate on CD. The archive consists of about 100 CDs. From this archive I selected 7000 images for compression using jpeg for the final published database. 225 conditions are included and these are listed alphabetically together with their alternative names (*synonyms*) in the *condition list*. Each image in the *image list* is associated with a radiological report, which is a list of findings identified by me on the image. For each condition there is a full list of radiological findings in the *findings list*, taken from all the findings associated with the images of the condition under consideration. The majority of conditions have images, which have been selected from several patients. A number in brackets follows each finding in the findings list. This represents the number of times the finding has been identified and gives some indication of how common a particular finding is in an individual condition. Over 22000 findings are incorporated in REAMS. In addition to the findings, every image is associated with some standard information. This includes the diagnosis, age range, gender, radiographic view (body part, e.g. upper limb) and projection (antero-posterior, lateral). The age ranges are divided into diagnostically significant groupings – fetus, neonate, infant, young child, older child, adolescent, adult. There is also a unique patient number (from the patient database) and an individual image number. When viewing images, in addition to the image findings, relevant clinical information about the patient is included when this is available. Each condition is also linked with an abstract about the condition and a comprehensive list of references. These have been duplicated from LDDB.

In summary, the database incorporates 7000 selected radiographic images and their findings, covering 225 conditions with their abstracts and reference lists.

The use of REAMS as a reference resource.

The electronic functionality allows the database to be used in several ways. Firstly it may be used as a standard reference atlas or textbook by selecting a specified condition for consideration. **Conditions** on the toolbar should be selected for this application. The condition is highlighted in the **conditions list** which is the page which opens by default. The condition can also be selected by typing in the first few letters of a diagnosis in the **find** box. The **literature details** can be selected to view the **abstract**. From the **abstract** page the **references, radiological findings** and **synonyms** can be selected.

Closing this page returns the user to the **conditions list** from which the user may select **images**.

The images are displayed as thumbnails. By default they appear in a **grid** format with 8/12 on each page and can be scrolled through using the scroll bar. Alternatively the thumbnails can be viewed in a **list** with about 4 on each page, each image associated with its list of radiological findings.

The head of the page confirms the diagnosis and states the total number of images for the condition and the number of patients from which they have been taken. Each thumbnail image has a heading with the diagnosis, projection, body part, age range, gender and unique patient number and image letter. The thumbnails and the order of images can be **sorted by** various parameters. These include **sorting by** view (body part), projection (AP, lateral), age range, or patient. This sorted list can be further **sorted by** the same parameters. For example the images could first be sorted by body part and then by age. This would enable the user to compare, for example, all views of the hands of a condition at increasing ages.
The full size view of the image may be seen by either clicking on the thumbnail image or by selecting **view image**. The selected image is displayed on the left side of the page with the findings for the image on the right although other positions can be chosen. Some images have arrows on them helping to identify unusual or subtle findings. The arrow on the image is associated with a red number, and this is seen also adjacent to the relevant finding describing the abnormality in the image finding list. Clicking on the arrow highlights the arrow and also the finding it indicates. Similarly, clicking on the numbered finding highlights the finding and the appropriate arrow on the image. If the user wishes to see the image without arrows they can be deselected (tick box **show arrows**). The image can occupy the entire screen by selecting **full screen** and parts can be further magnified by clicking the mouse cursor (magnifying glass) over the area of interest on the image. Return to normal magnification is achieved by clicking the right mouse button. **Restore** returns to the previous page. Also from this page, notes about the image (**image notes**) and notes about the clinical details of the patient (**patient notes**) can be found and the user can also view the abstract about the condition. Using the forward and backward arrows, the user can step through the images from the sorted thumbnail list. **Closing** this page returns to the thumbnail page.

From the *conditions list* or *images* pages the *findings* associated with the images of the condition can be displayed in either *full* reporting form (flattened epiphyses of the humeri x1 and the epiphyses of the proximal humeri are markedly flattened x1 and flattened epiphyses of the proximal humeri x1), *standard* form (flattened epiphyses of the humeri x1 and flattened epiphyses of the proximal humeri x2) or *short* form (flattened epiphyses of the humeri x3). The numbers after the findings indicates the number of images with that finding for this condition. The findings are grouped by *film* under the headings of the body parts (skull, chest, spine) or by *abnormality* (abnormal size, abnormal density, abnormal development, abnormal shape) or by *set* (long bones, epiphyses, metaphyses). One or more findings can be highlighted or chosen using the *select* buttons and *view images of selected* retrieves images in the thumbnail format demonstrating the selected finding(s). The *conditions* tab returns the user to the findings page where the highlighted findings can be deselected and others chosen.

The use of REAMS as a diagnostic aid

This task is performed using **features** on the menu bar. This is the page used to select and search on findings, which have been identified from an unknown case.

Features are shown on the Feature Search page and are organized into two parts: **body parts** and **abnormalities** or findings.

Body parts are listed in a tree structure, which is shown in the top right box of the **Feature search** page. The tree shows a hierarchical list of body parts. A body part has other body parts underneath it, which are part of that body part. For example upper limbs is a top-level body part and it has other body parts (forearms, hands, fingers) underneath it because these are all part of the upper limbs. There is no limit to the number of levels the tree can have; there are different numbers of levels in different parts of the tree. This structure means that a high level of anatomical precision can be achieved.
For example, when describing an abnormality of the epiphyses of the proximal phalanges of the index fingers, the sequence in the tree structure would be:-
upper limb

hands

fingers

index fingers

phalanges

proximal phalanges

epiphyses

A body part can be expanded, to show all the body parts underneath it, or collapsed, to hide them again, in any of the following ways:

- Click on the plus or minus sign to the left of the feature
- Double click on body part's name
- Use the plus (+) and minus (-) keys on the keyboard. Pressing plus while a body part is selected will expand it and pressing minus will collapse it.
- Use the left and right arrow keys on the keyboard. Pressing right while a body part is selected will expand it and pressing left will collapse it.

If a body part does not have a plus or minus sign to the left of it then there are no body parts underneath it. When a body part is expanded any other expanded body part on the same level will be collapsed automatically.

All body parts in the tree have a number next to them in square brackets. This is the number of findings there are for that body part specifically, as well as for all the body parts underneath it in the tree. It does not imply the number of images for this body part because an image can include more than one finding.

Abnormalities are shown in the box in the lower right side of the page, below the **body parts**. If a body part in the tree is selected then a list of abnormalities for that body part will be shown here.

Each abnormality has a number next to it in square brackets specifying the number of findings for that abnormality.

To find a feature quickly, enter it in the **Find** box, which is between the body parts and abnormalities. It can be entered in the **Find** box by either typing it or selecting it from the list. When the finding is entered, select either the Enter key on the keyboard or the **Find** button. If there is only one feature with the name specified then the tree will be expanded to the correct place and the relevant body part and abnormality (if an abnormality was specified) will be selected. If there is more than one feature found then they will be displayed as a list in the **Find** window.

To select a feature in the list, click on it. To select a range of features, click and hold the left mouse button on the first feature, then move the mouse to the last one and release the button. To select features that are not next to each other, hold down the Control key while clicking on the features to be selected with the left mouse button.

To add a body part or an abnormality to a **criterion**, first select the **criterion** box.
A body part can be added by selecting it in the tree and clicking the **Add** button or an abnormality can be added. A range of multiple abnormalities can be selected and added at once.
Features can also be added to a **criterion** using drag and drop. When a feature is dropped in a **criterion** that criterion is automatically selected. Multiple abnormalities can be added at once by using drag and drop with the group of selected abnormalities.

Up to nine criteria can be specified for one search. The **criteria** boxes are shown on the left of the **Feature search** page.
Features within the same **criterion** are alternatives, which means that if an image only has one of the features in a **criterion** then it matches. Checking the **mandatory** box for a **criterion** means that an image or condition must match that **criterion**. Each **criterion** also has a **Broad** checkbox. This specifies whether the search should be broad or narrow.
When doing a narrow search any feature searched for will include the findings for that feature itself and the findings for all features underneath it in the feature tree only. A broad search does a more complicated search of other features on the same or a higher level than that feature, to find other features that are also related to it.

For example, phalanges of the fingers is under fingers in the tree. Phalanges of the fingers means abnormalities of all (or most) of the phalanges. A narrow search on phalanges of the fingers will simply return all the findings underneath this feature, therefore returning only findings that are abnormalities of all (or most) of the phalanges. A broad search, on the other hand, will do further searches on the same and higher levels to find abnormalities of one or more of the phalanges. So it will also find phalanges of the index fingers, which is under index fingers, which is in turn under fingers in the tree. It will also find phalanges of the middle fingers, ring fingers, and little fingers.

Narrow searches are only used to retrieve images of precisely defined specific abnormalities. A user needs to have some experience in the field of skeletal dysplasias when using narrow searches. A broad search is generally used during a consultation, bringing forward more, less precise but related, images and conditions for consideration.
The number next to the body parts in the tree is the number of findings that would be found when doing a narrow search only. A broad search could return many more.

As well as specifying features that an image should have, some filters can also be applied using the. **Filters** button to specify retrieving images only from defined ages or age ranges or from keywords used in the free text in the **image notes** or **patient notes**. The ages that can be selected are fetus, neonate, infant, young child, older child, adolescent, and adult. To set a range of ages, choose a different age for each of the **Age Range** boxes or just a single age can be selected.
Words entered in the **Keywords text** box need to be in either the **image notes** or the **patient notes** in order for the image to be found. Multiple words can be entered by separating them with a space. If multiple words are entered then they all need to appear in the image or patient notes (although not necessarily together) in order for an image to be found.

The number of criteria an image or condition needs to match is entered just below the list of criteria.
It specifies how many of the criteria need to be matched in order for an image to be included.
Match criteria by specifies one of two methods that should be used to match the criteria.

Matching by **Image** is the simplest method. This means that an image will be included in the results if it shows at least one finding in the number of criteria specified.
Matching by **Condition** is used when searching for conditions with groups of features. When this option is selected the results of the search are images whose condition or diagnosis satisfies the minimum search criteria. In addition each image must match at least one criterion. Matching this way is appropriate when REAMS is being used to assist in making a diagnosis.

When the findings have been entered in the **criterion** boxes, the **mandatory** and **broad** boxes have been set, the **filters** applied and the **number of criteria to match** decided, the **Search** button initiates the search.

If no images are found, a message will be displayed, otherwise the **Images** page will show all the images that were found, displayed as thumbnails in a grid format. When the images are viewed in a list or individually with the image findings, the findings that match the criteria from the search are highlighted.

The thumbnail images can be sorted according to the criteria already identified, but an additional criterion is to **sort by condition**. This is particularly useful when images from several diagnoses are being displayed.

At the top of the **Images** page is the statistics bar, which shows the number of images, the number of patients, and the number of conditions in the list from the images which have been retrieved. The list of diagnoses can be seen by placing the cursor over **Conditions** for a short time. If the list is long, there will be a scroll bar on the right-hand side. The list will disappear as soon as the cursor is removed from either the list or **Conditions**. More detailed statistics for the conditions are available in **Condition Statistics** at the right of the statistics bar. This lists all the conditions in the image list, and also the number of images and patients displayed with each condition.

The **Clear Criterion** button removes all the findings from a single criterion.
The **Clear Search** button clears all the criteria and filters.

Please see appendix C for a sample diagnostic consultation using REAMS

Chapter 8. Clinical trial: comparison of the Image Database (a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias – REAMS) with Standard methods of diagnosing Skeletal Dysplasias by General Radiologists.

Aim

The purpose of this clinical trial was to test the hypothesis that the use of the database of images (a Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias, REAMS) could improve the diagnostic accuracy of general radiologists in this field.

Background

The retrospective evaluation of diagnostic accuracy achieved by different clinical specialties has identified that general radiologists only identify the correct diagnosis in 31% of cases compared to an accuracy of 57% by clinical geneticists.

The clinical trial comparing the use of a knowledge-based expert system with standard methods of diagnosis by general radiologists achieved a dramatic improvement in accuracy. The system arrived at the correct diagnosis in 72% of cases compared with 50% when using textbooks. The correct diagnosis was included in the top three diagnoses (identified as a 'good patient outcome') in 86% of cases using the expert system compared with 53% using textbooks.

Several picture archive (image) databases are available to aid diagnosis in this field. LDDB (The London Dysmorphology Database) [19] and POSSUM (Pictures of Standard Syndromes and Undiagnosed Malformations) [17], provide powerful search facilities and a comprehensive clinical approach, largely illustrating dysmorphic features. OSSUM is a radiographic database of skeletal dysplasias linked to POSSUM, and therefore provides an equivalent resource to REAMS.

OSSUM has been tested in a clinical setting [24]. This trial concluded that an accurate diagnosis was achieved in exactly the same proportion of tests using the database compared with those using standard methods of diagnosis.

8.1 Method

Exactly the same trial protocol and material was used as in the trial of the expert system (pages 64-67). The only differences were the use of REAMS rather than the expert system and the eight radiologists participating in the trial. A new group of radiologists was selected using the same criteria. They had recently completed their final UK FRCR examination and therefore had received recent teaching on the subject, read the most recent literature and studied some texts on skeletal dysplasias in their examination preparations. The new group of radiologists was selected to avoid a possible learning process from both the use of the expert system and repeat viewing of the same case material. The radiologists participating in the trial had no experience of using REAMS before the trial. They were each given a ten-minute introduction to using the database and the instruction manual.

8.2 Results

Table XVIII
Table of results

Diagnosis	Correct ranked 1st	Correct suggested	Unknown	Incorrect suggested	Incorrect concluded
REAMS	69%	5%	3%	3%	21%
Books 2	38%	4%	4%	31%	23%

Table XIX
Assessment of patient outcome

Outcome	Good outcome - Correct in top 3 diagnoses	Poor outcome – Incorrect diagnosis suggested or concluded
REAMS	74%	24%
Books 2	42%	54%

Correct diagnosis identified or suggested by REAMS = 74%
Correct diagnosis identified or suggested by books = 42%

Incorrect diagnosis identified or suggested by REAMS = 24%
Incorrect diagnosis identified or suggested by books = 54%

The results can be divided into two categories of outcome:
Favourable, where the correct diagnosis is either concluded or ranked second or third
Unfavourable, where an incorrect diagnosis is either suggested (ranked second or third) or concluded. In other words, the correct diagnosis is not in the top three suggested diagnoses.

As far as the patient is concerned, a favourable outcome enables an accurate diagnosis to be made, sometimes following further targeted specific diagnostic tests such as clinical and dysmorphic evaluation or biochemical or molecular studies. A firm

diagnosis means that meaningful genetic counselling for the patient and their family can be offered. With an understanding of the natural history of the condition, complications can be predicted and sometimes prevented and management planned. In some conditions treatment may even be curative.
An error in diagnosis is an extremely unfavourable outcome for the patient and leads to inaccurate labelling and inappropriate management.

Overall, REAMS produced 74% favourable outcomes; using the books resulted in 42% favourable outcomes. We can therefore conclude that using the system significantly increased the likelihood of a general radiologist reaching a favourable outcome.

Overall, REAMS produced 24% errors resulting in an unfavourable outcome, as compared to 54% errors produced when using textbooks. The radiologists using REAMS had no prior experience of using the multiple functions of the database and had only a short introduction. Increased familiarity with the capabilities of the database could potentially further enhance diagnostic accuracy.

8.3 Discussion and Conclusions

The results of the clinical trial of OSSUM were significantly different from those using REAMS. There are several possible reasons for this. Firstly the aims of the trials were different. The OSSUM trial was assessing the usefulness of OSSUM in improving the diagnostic accuracy of a group of paediatric radiologists in a single centre in the USA. The REAMS trial was conducted in the UK where there is only a handful of centres with more than three paediatric radiologists, and it was being assessed for improved diagnosis of general radiologists with minimal paediatric radiology experience. Secondly, the trial protocol was different. In the OSSUM trial, the four radiologists testing the database were all paediatric radiologists with an average of 9.5 years experience. In the REAMS trial the eight radiologists had a maximum of 6 months experience in paediatric radiology. This means that the background knowledge of paediatric normal variants and normal development and ossification of the skeleton from the fetus through to adulthood, was different in the two groups. In the trial of OSSUM only 20 cases were tested using four radiologists compared to 32 cases with eight radiologists using REAMS. The sample size was small in the OSSUM trial.

The use of REAMS significantly enhanced the diagnostic accuracy of general radiologists in the field of skeletal dysplasias and potentially improves patient outcome and quality of life.

Chapter 9. Combined results from the clinical trials of the two computer based diagnostic systems

Although the clinical trials of the expert system and of REAMS were conducted on separate occasions it is possible to make some comparisons between the results of the two systems. This is because the same material was used (the same 32 skeletal surveys) and exactly the same method was employed using eight comparably qualified general radiologists for each trial. Each radiologist evaluated eight skeletal surveys using standard methods of diagnosis (books) and eight skeletal surveys using either the expert system or REAMS. **Table XX** shows the combined results from the two clinical trials. Using the standard, book method for diagnosis, although the second group of radiologists achieved fewer correct diagnoses compared with the first group, and suggested relatively more incorrect diagnoses, this was not a significant difference. The results using the standard book method for diagnosis were therefore combined to give an evaluation from all sixteen radiologists and compared to the results from the expert system and REAMS in **Table XXII.**

Table XX
Combined results from the two clinical trials

Diagnosis	Correct ranked 1st	Correct suggested	Unknown	Incorrect suggested	Incorrect concluded
Expert system	72%	14%	0%	9%	5%
Books 1	50%	3%	2%	19%	27%
REAMS	69%	5%	3%	3%	21%
Books 2	38%	4%	4%	31%	23%

Table XXII
Results from the two clinical trials combining the results using books

Diagnosis	Correct ranked 1st	Correct suggested	Unknown	Incorrect suggested	Incorrect concluded
Expert System	72%	14%	0%	9%	5%
REAMS	69%	5%	3%	3%	21%
Books	45%	2%	3%	24%	25%

Both the expert system and REAMS arrived at the correct diagnosis in significantly more cases (about 70%) than when the books were used (45%). However the correct diagnosis was suggested (ranked second or third) in 14% using the expert system compared to only 5% using REAMS and 2% using the books. This difference between arriving at a correct diagnosis and including the correct diagnosis in second or third position when using REAMS is likely to be the result of the ability to match images. An incorrect diagnosis was made significantly more frequently using REAMS (21%) compared to the expert system (5%). It is most likely that this is related to evaluation of secondary triggers by the expert system. In this process, close differential diagnoses with some similar radiological features are explored, but may be excluded on the basis of other specific findings on direct questioning.

An assessment of patient outcome can be made using the results from the two clinical trials and also by combining the results of all sixteen radiologists using standard book diagnosis - **Table XXI** and **Table XXIII**.

Table XXI Assessment of patient outcome from the two clinical trials

Outcome	Good outcome - Correct in top 3 diagnoses	Poor outcome – Incorrect diagnosis suggested or concluded
Expert system	86%	14%
REAMS	74%	24%
Books 1	53%	45%
Books 2	42%	54%

Table XXIII
Assessment of patient outcome from the two clinical trials combining the results using books.

Outcome	Good Outcome – Correct in top 3 diagnoses	Poor Outcome – Incorrect diagnosis suggested or concluded
Expert system	**86%**	**14%**
REAMS	**74%**	**24%**
Books	**47%**	**49%**

In clinical practice the diagnoses ranked in the first three positions will be evaluated in more detail to determine the correct diagnosis. A correct diagnosis falling in the first three ranked positions constitutes a good patient outcome. Conversely, if the correct diagnosis is not included in the top three positions, which means that incorrect diagnoses have been suggested or concluded, this constitutes a poor or adverse patient outcome. The expert system achieves a good patient outcome in 86% of cases, compared to 74% using REAMS and only 47% using books. The number of patients with a good outcome (47%) is almost the same as with a poor outcome (49%) when standard diagnosis using books is used.

In both clinical trials the number of times an 'unknown' diagnosis was made was strikingly low. In clinical practice about 50% of skeletal surveys referred for diagnosis of a suspected skeletal dysplasia are unknown (**Table I page 19**). The most likely explanation for this low number is that all the participating radiologists were told that the skeletal surveys were drawn from patients with confirmed diagnoses within the group of sclerosing and cranio-tubular disorders. This bias was therefore built into the trial design. If a low number of 'unknown' diagnoses is still achieved with clinical usage of either of the systems, it may indicate inappropriate false confidence as a result of using computerised aids to diagnosis. There would need to be some awareness of this by the user.

In conclusion, both computer aids to diagnosis of skeletal dysplasias (the expert system and REAMS) result in a significant improvement in diagnostic accuracy of general radiologists. It is possible that more experience with the functionality of REAMS would result in further improved diagnostic accuracy.

The final aim is to link the expert system with the electronic database of images. It is anticipated that this would further enhance diagnostic accuracy. The clinical aim should be to identify an appropriate limited number of diagnostic possibilities with a view to recommending targeted sets of biochemical and genetic mutation screens to confirm the final diagnosis. This approach is currently used by the expert panel of the European Skeletal Dysplasia Network (ESDN), currently with input from only one radiologist. The routine use of computer aided systems to improve radiological diagnostic possibilities should improve overall diagnostic accuracy.

References

1. Hall CM, International Nosology and Classification of Constitutional Disorders of Bone (2001). Am J Med Genet 2002, 113:65-77.

2. Poznanski AK. Bone dysplasia symposium introduction. Pediatr Radiol 1994;24:383.

3. Stoll C, Dott B, Roth M-P Alembik Y. Birth prevalence rates of skeletal dysplasias. Clin Genet 1989;35:88

4. Gormley J, Wynne-Davies R. The prevalence of skeletal dysplasias. JBJS 1985;67B:133-137

5. StevensJM, Kendall BE, Crockard HA, Ransford A. The odontoid process in Morquio-Brailsford's disease. The effects of occipitocervical fusion. L Bone Joint Surg Br 1991;73:851-858.

6. Keiper GL Jr, Koch B, Crone KR. Achondroplasia and cervicomedullary compression: prospective evaluation and surgical treatment. Pediatr Neurosurg 1999;31:78-83.

7. Al-Mefty O, Fox JL, Al-Rodhan N Dew JH. Optic nerve decompression in osteopetrosis. J Neurosurg 1988;68:80-84.

8.Cheow HK, Steward CG, Grier DJ. Imaging of malignant infantile osteopetrosis before and after bone marrow transplantation. Pediatr Radiol 2001;31:869-875.

9. McKusick VA. 'Mendelian Inheritance in Man. Catalogs of Autosomal Dominant, Autosomal Recessive and X-linked Phenotypes' 8th edition. 1988 Baltimore and London: The Johns Hopkins University Press.

10. Spranger JW, Brill PW, Poznanski A. Bone dysplasias: an atlas of genetic disorders of skeletal development. 2nd edition. 2002 Oxford University Press, Oxford

11. Taybi H, Lachman RS. Radiology of Syndromes, Metabolic disorders and Skeletal Dysplasias. 4th edition. 1996 Year Book Medical Publishers, Littleton, Mass

12. Wynne-Davies R, Hall CM, Apley A. An Atlas of Skeletal Dysplasias. Churchill Livingstone, Edinburgh 198513.

13. DiLiberti JH. Use of computers in dysmorphology. J Med Genet 1988 25:445-453.

14. Kozlowski K, Beighton P. Gamut index of skeletal dysplasias: an aid to radiodiagnosis. 1984 Springer, Berlin Heidelberg New York 1984

15. Hall CM, Washbrook J. A Radiological Electronic Atlas of Malformation Syndromes and Skeletal Dysplasias. Oxford University Press, Oxford.

16. Bankier A, Keith CG. POSSUM: the microcomputer laser-videodisk syndrome information system. Opthal Paediatr Genet 1989;10:51-52.

17. Bankier A, Marquet J. Possum – a computer laser videodisk system for syndrome diagnosis and learning about diagnosis. Am J Hum Genet 1991;49(4):28.

18. Bankier A, McGill JJ, Danks JJ, McGill JA, Danks DM. Dysmorphology: problems in nomenclature. Dysmorph Clin Genet 1988;2:25-50.

19. Winter RM, Baraitser M. The London Dysmorphology Database: A Computerised Database for the Diagnosis of Rare Dysmorphic Syndromes. 6th Edition. 2001 Oxford University Press, Oxford.

20. Winter RM, Baraitser M, Douglas JM. A computer database for the diagnosis of rare dysmorphic syndromes. J Med Genet 1984;21:121-123.

21. Keravnou ET, Dams F, Washbrook J, Hall CM, Dawood R, Shaw D. Modelling diagnostic skills in the domain of skeletal dysplasias. Computer Methods and Programs in Biomedicine 1994;45:239-260. **Appendix E**

22. Evans CD. Computer systems in dysmorphology. Clin Dysmorphology 1995;4:185-201.

23. Guest SS, Evans CD, Winter RM. The Online London Dysmorphology Database. Genetics in Medicine 1999;1:207-212.

24. Harned RK, Patrick LE, Gay BB, Atkinson GO, Niemer PK, Wyly JB, Clark WS. Standard method of diagnosis versus use of a computer database in the evaluation of skeletal dysplasias. Pediatr Radiol 1996;26:887-890.

25. Weiner F, Anneren G. PC-based system for classifying dysmorphic syndromes in children. Comput. Methods Programs Biomed 1989;28(2):111-117.

26. Yamamoto K, Sudo M, Shigematsu Y, Fukui M, Masukawa. The development of a personal computer-based medical consultation system for the diagnosis of congenital malformation syndromes using MUMPS. Med Inform 1990;15(4):355-362.

27. Dawood RM, Hall CM, Shaw DG, Keravnou ET, Washbrook J. A diagnostic Expert System for Skeletal Dysplasias and Syndromes: an Almost Ideal Application for Expert Computer Systems in Medicine. (Proc of the International Congress of Radiology, Paris 1989).

28. Dawood RM, Hall CM, Shaw DG, Keravnou ET, Washbrook J, Dams F. Model-based expert system for diagnosis in paediatric skeletal radiology. Radiology 1990;177:208.

29. Hall CM, Shaw DG. Computer diagnosis of skeletal dysplasias and malformation syndromes. Acta Paediatr Suppl 1994;406:73-6. **Appendix E**

30. Keravnou ET, Dams F, Washbrook J, Dawood RM, Hall CM, Shaw D. Background knowledge in diagnosis. Artif Intell Med 1992;4:263-279. **Appendix E**

31. Keravnou ET, Johnson L. Intelligent handling of data by integration of commonsense reasoning. Knowledge-Based Systems Journal 1987;1:32-42.

32. Keravnou ET, Washbrook J, Dawood RM, Hall CM, Shaw DG. A model-based diagnostic expert system for skeletal dysplasias. Proc AIME 89, Hunter J, Cookson J, Wyatt J, Eds. 47-56 Springer-Verlag, Berlin 1989. **Appendix E**

33. Keravnou ET. (Ed) Medical temporal reasoning, Special issue of Artif Intell Med1991;3(6)

34. Keravnou ET, Washbrook J. A temporal reasoning framework used in the diagnosis of skeletal dysplasias. Artif Intell Med 1990;2:239-265.

35. Giedion A. The weight of the fourth dimension for the diagnosis of genetic disease. Pediatr Radiol 1994;24(6):387-91.

36. Wynne-Davies R, Hall CM. Two clinical variants of spondyloepiphyseal dysplasia congenita. JBJS 1982;64B:435-441.

Appendix A

Example of a dysplasia frame

Name: Morquio disease

Synonyms: Mucopolysaccharidosis type IV, MPS IV.

Clinical features: trunk short
sternal protrusion
+- scoliosis
knock-knees
ligamentous laxity
corneal opacities
IQ normal
AR
deafness progressive
dentition poor
presents from 1 year

Biochemical features:
keratin-sulphate in urine excess

Radiological features:

Skull mandibular condyles concave
mastoid air cells underdeveloped

Spine odontoid hypoplasia
cervical spine instability
platyspondyly throughout
lumbar vertebral bodies anterior tongues
vertebral bodies posterior scalloping
thoraco-lumbar kyphosis from 4 years

Pelvis iliac wings flared
acetabula sloping
iliac bases hypoplasia

Limbs coxa valga
femoral capital epiphyses progressive resorption
femoral capital epiphyses flat from 2-10 years

	femoral capital epiphyses absent ossification from 10 years
	genu valgum
	+- diaphyses slight widening
	metaphyses flared, irregular
	epiphyses small, irregular
	distal radius metaphysis sloping
Hands	2-5 metacarpals proximal pointing
	carpal bones small, irregular
Chest	ribs posterior constriction
	ribs anterior widening
	thorax short
	chest AP diameter increased
	pectus carinatum
	+- clavicles broad
	+- heart enlarged
Typical features	keratin sulphate in urine excess
Sufficient set	platyspondyly throughout
	odontoid hypoplasia
	ribs anterior widening
	acetabula sloping
	iliac bases hypoplasia
	femoral capital epiphyses abnormal from 2 years
	epiphyses small irregular
	2-5 metacarpals proximal pointing
	metaphyses flared, irregular, sloping
Exclude if	acetabula horizontal
Primary triggers	platyspondyly
	vertebral bodies anterior tongues
	epiphyses small, irregular
	ribs anterior widening
	acetabula sloping
	femoral capital epiphyses progressive disappearance
	metacarpals proximal pointing
	corneal opacities

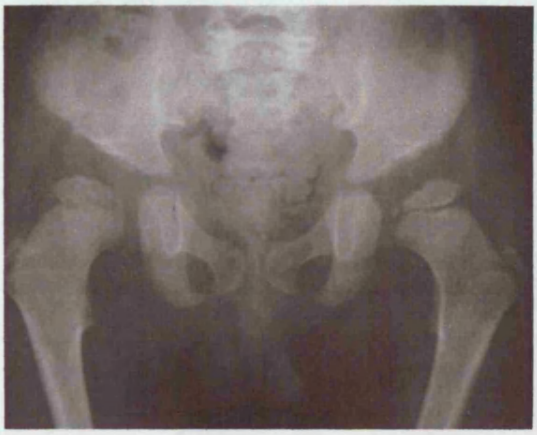
	ligamentous laxity	
	metaphyses irregular, flared, sloping	
	femoral capital epiphyses absent ossification	
	keratin sulphate in urine excess	
Secondary triggers	presents from birth	SEDC
		metatropic dysplasia
		diastrophic dysplasia
		Kniest disease
		fibrochondrogenesis
	iliac wings not flared	
	acetabula flat	
	ribs no anterior widening	pseudoachondroplasia
	joints limited mobility	
	iliac wings lace-like	
	hand epiphyses cone-shaped	
	vertebral end plates central notches	DMC
	2-5 metacarpals no proximal pointing	
	ribs normal	
	vertebral bodies no central tongues	SMD
	myopia	
	Cleft palate	
	stature mildly reduced	
	epiphyses mildly abnormal	Stickler's syndrome

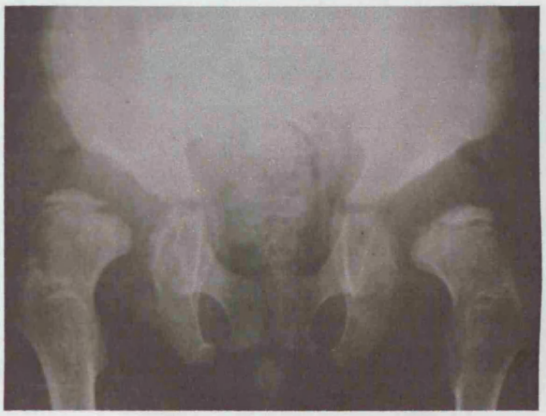
Appendix B. Sample consultation with the expert system

This illustrates most aspects of the diagnostic expert system.

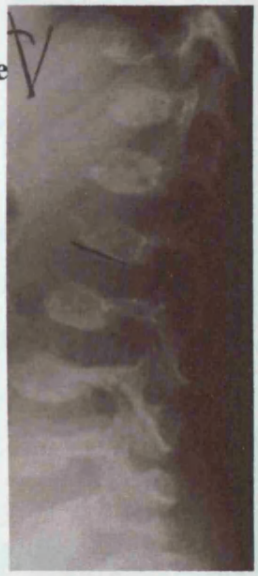
The radiographs available in this consultation included:

Pelvis at the ages of 2 and 9 years.

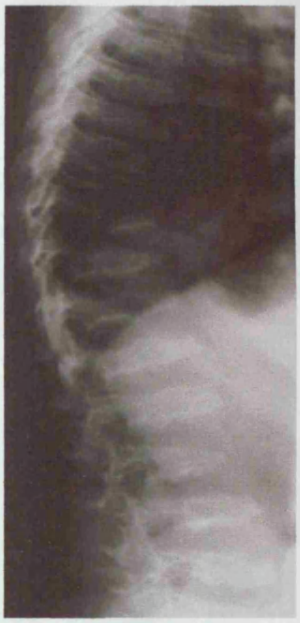




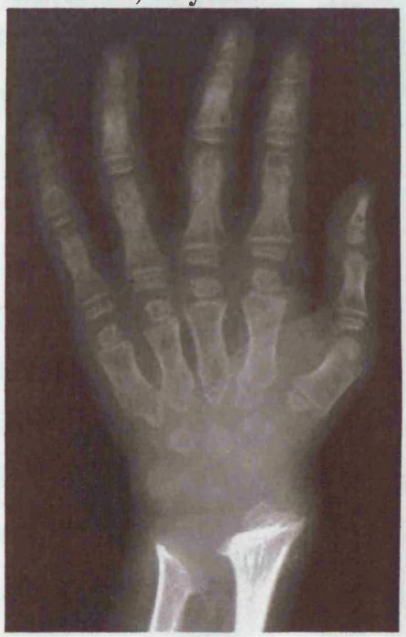
Lateral spine
2 years



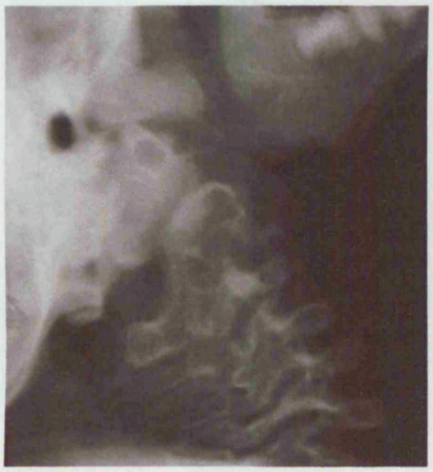
Lateral spine
7 years



Left hand, 10 years



Lateral cervical spine, 10 years



Enter patient age (yrs mths), perinatally lethal, stillborn, prenatal: (10 0)
Is patient alive (y/n)? y
Enter patient sex (m,f,unk) m
Consanguinity (y,n,unk) unk
Enter age of patient when the problem first presented (yrs mths or unk) unk
Enter youngest age of patient where there are radiographs showing some abnormality (yrs mths) 2 0

The age of the patient at first presentation may well not be known, but the user can still identify an upper limit of presentation from the age at the earliest films. This may be the current age of the patient.

Collecting clinical, histological and biochemical findings
Enter each finding on a separate line. Terminate list with E
Short-trunk dwarfism
E
Collecting radiological findings
Enter each finding on a separate line. Terminate list with E
platyspondyly entire spine
vertebral end plates irregular
For each of the following queried findings enter y,n or unk, or enter correct temporal aspect:
vertebral end plates central notches? n

This is an example of a prompting question. The irregularity of the vertebral end plates may be due to the presence of central notches. The user continues to enter radiological findings.

lumbar spine anterior beaking
odontoid peg absent
lumbar lordosis prominent
sacral angulation
femoral capital epiphyses flat irregular (at yrs 9)
femoral necks wide

coxa valga
iliac wings wide
iliac crests irregular
for each of the following queried findings enter y,n or unk, or enter correct temporal aspect:
iliac crests lace-like? n

This is another example of a prompting question.

2^{nd} – 5^{th} metacarpals proximal pointing
epiphyses irregular
carpal bones small irregular
ulnae short
E

The user has now answered the general questions raised by the system and has entered the initial clinical and radiological case findings. Each item of information is processed to determine if it is a hard finding and to decide whether possible hypotheses should be triggered. The system has identified the following as hard abnormalities – short-trunk dwarfism, platyspondyly, irregular vertebral end plates, wide femoral necks, wide iliac wings, metacarpals which are short and point proximally and irregular epiphyses. At this stage possibilities are generated if one of their primary triggers matches. Possibilities are excluded if their expected presentation time does not match that of the patient, or if one of the exclude-if features match or if some other exclusion criterion is met. The system lists each possibility with its reason for inclusion or exclusion.

excluding DMC. Reason: exclude-if match

For DMC a significant feature is that of lace-like iliac wings between the ages of 3 and 15 years old. The system has been told that the patient does not have lace-like iliac wings (pelvis aged 9 years).

including metatropic dysplasia. Reason: primary trigger match
including SMD. Reason: primary trigger match

including Kniest dysplasia. Reason: primary trigger match

These 3 possibilities are generated because they all have platyspondyly as a primary trigger. Another primary trigger for Kniest dysplasia, which also matches, is wide femoral necks. Morquio disease is suggested by a primary trigger (platyspondyly), but since it has an exclude-if feature (acetabula horizontal), the system raises a query.

for each of the following queried findings enter y,n or unk, or enter correct temporal aspect:
acetabula horizontal ? n
including Morquio disease. Reason: primary trigger match
including SEDC. Reason: primary trigger match
excluding Stickler syndrome. Reason: exclude if match

Stickler syndrome has localised platyspondyly. This patient has generalised platyspondyly excluding this diagnostic possibility.

including pseudoachondroplasia. Reason: primary trigger match
including diastrophic dysplasia. Reason: primary trigger match

Again platyspondyly is the case finding which has triggered these possibilities. This illustrates that one primary trigger is often associated with more than one dysplasia. In addition pseudoachondroplasia is triggered by the finding metacarpals proximal pointing.

excluding achondrogenesis type I. Reason: lethal, but case born and alive
exclude brachyolmia. Reason: dysplasia presents after 5yrs

Although the age of the patient at which the problem first presented is not known, the system has been told that radiographs taken at the age of 2 years show abnormalities. The patient was therefore presenting with problems at least as early as 2 years. On the basis of this information, brachyolmia is excluded on the basis of presenting after about the age of 5 years. Also the system has been told that the epiphyses are irregular and the carpal bones are small which implies that the epiphyses are small. The system knows

that the epiphyses are abnormal and brachyolmia is also excluded on the basis of the expectation of normal epiphyses in brachyolmia.

excluding atelosteogenesis I. Reason: lethal but case born and alive

The triggered possibilities are ranked from the perspectives of the explanatory power (how well they account for all the case findings) and sufficient set matches (how well their sufficient features group are satisfied). In each ranking the typical match (whether a typical feature for the dysplasia matches) is used as the primary criterion since such matches categorically conclude the given dysplasia.

*Initial context from the **explanatory power** perspective*

Possibility	*Typical match?*	*Explanatory power*
Morquio disease	*0*	*60*
pseudoachondroplasia	*0*	*36*
Kniest dysplasia	*0*	*31*
metatropic dysplasia	*0*	*30*
SMD	*0*	*25*
diastrophic dysplasia	*0*	*21*
SEDC	*0*	*15*

*Initial context from the **sufficient set** match perspective*

Possibility	*Typical match?*	*%Sufficient set match*
SMD	*0*	*50*
Morquio disease	*0*	*43*
pseudoachondroplasia	*0*	*30*
Kniest disease	*0*	*27*
metatropic dysplasia	*0*	*25*
diastrophic dysplasia	*0*	*20*
SEDC	*0*	*0*

In this consultation no typical matches have occurred (if more than one typical match occurs simultaneously then, on the assumption that both the system knowledge and the user input are valid, the patient exhibits multiple dysplasias or a new condition).

On the basis of the above we focus on the following differential:

Possibility	*Typical match*	*Explanatory power*	*%Sufficient set match*
Morquio disease	*0*	*60*	*43*
SMD	*0*	*25*	*50*

Although 7 possibilities were triggered the system focuses on 2 of them only; the others are temporarily removed from consideration. The selected possibilities are collectively explored. The system asks non-leading questions drawn from the unknown sufficient and typical features of these possibilities.

for each of the following queried findings enter y,n or unk, or enter correct temporal aspect:
IQ normal? unk
hearing impaired? unk
corneal opacities? unk
joints limited mobility? unk
gait waddling? unk

The user can enter a question with 'unknown' even when the available radiographs provide an answer, if the user is unsure of the answer.

femoral capital epiphyses abnormal from yrs 2? from 2 to 9 yrs

The system knows that the femoral capital epiphyses are abnormal at the age of 9 years, but it needs to know if they have been abnormal from about the age of 2 years. The available radiographs show that the femoral capital epiphyses were abnormal at the ages of 2 and 9 years, and it is reasonable to assume that they had been abnormal between those ages. The user gives the information by entering the corresponding temporal aspect in response to the above question.

epiphyses almost normal? n
iliac bases hypoplastic? y
iliac bones short? unk

ribs anterior widening? unk

Although there are two views of the pelvis, the user is unable to decide whether the iliac bones are short or not. However because there is no view of the chest, the question about ribs is truly unknown.

keratin sulphate in urine excess? unk
Promoting Morquio disease to a working hypothesis. Reason: sufficient hard match

At the end of the exploration of the initial differential Morquio disease is the only possibility which is promoted to a working hypothesis because it is the only one which adequately covers the identified hard abnormalities.

Current differential

	Typical match	*Explanatory power*	*%Sufficient set match*
Morquio disease	*0*	*68*	*66*
Strong differential			
Morquio disease	*0*	*68*	*66*

Differentiating through secondary triggers the cluster of hypotheses
Morquio against SMD, pseudoachondroplasia, Kniest disease, metatropic dysplasia

Since Morquio disease is the only working hypothesis it forms the strongest diagnostic possibility. The above numbers indicate that there is currently no typical match (the typical feature is an excess of keratin sulphate in the urine, which is unknown in this case), that 68% of the sufficient feature set is satisfied and that Morquio disease explains 66% of the entire set of case findings. On this basis Morquio disease is considered a strong hypothesis but it is not considered sufficiently better than its closest competitors to justify concluding it.
At this stage a differentiation strategy is therefore applied. Morquio disease is the system's focus of attention and is differentiated against its closest competitors through secondary triggers (see page 25). SEDC and diastrophic dysplasia are possible alternative diagnoses but their explanatory power and sufficient set matches are

relatively low and although activated are not considered in the differentiation process. Two dysplasias are competitors if they have highly overlapping feature sets.

Considering secondary triggers in the context of Morquio disease
If presents from birth
then consider Kniest disease
metatropic dysplasia
result all findings unknown
If iliac wings not flared
acetabula horizontal
no anterior widening
then consider pseudoachondroplasia
result rule has been refuted because
iliac wings not flared
acetabula horizontal
do not hold for the case
If 2-5 metacarpals no proximal pointing
ribs normal
vertebral bodies no central tongues
then consider SMD
result rule has been refuted because
2-5 metacarpals no proximal pointing
do not hold for the case
Considering secondary triggers in the context of SMD
If coxa valga
then consider Morquio disease
result 100% match
If epiphyses abnormal
then consider Morquio disease
result 100% match

The % match indicates the proportion of the secondary trigger antecedents that hold.

Considering secondary triggers in the context of pseudoachondroplasia

If *corneal opacities*

limbs short

trunk short

coxa valga

then consider Morquio disease

result 50% match

unknown findings are corneal opacities

limbs short

Considering secondary triggers in the context of Kniest dysplasia

If *cleft palate absent*

then consider Morquio disease

result all findings unknown

If *corneal opacities*

then consider Morquio disease

result all findings unknown

If *coxa valga*

then consider Morquio disease

result 100% match

If *femoral capital epiphyses progressive disappearance*

then consider Morquio disease

result all findings unknown

If *metacarpals proximal pointing*

then consider Morquio disease

result 100% match

Considering secondary triggers in the context of metatropic dysplasia

If *acetabula sloping*

coxa valga

Then consider Morquio disease

result 100% match

The system evaluates the relevant secondary triggers under each of the diagnostic possibilities. The antecedent of a secondary trigger either matches completely or

partially or is refuted. The system then asks the user about the unknown features of the partially matching secondary triggers.

For each of the following queried findings answer y,n or unk or enter correct temporal aspect
limbs short? unk
femoral capital epiphyses progressive disappearance? y
cleft palate? unk

In this case there are four such unknown features: limbs short, corneal opacities, cleft palate and progressive disappearance of the femoral capital epiphyses. The system has already been told that it is not known whether the patient has corneal opacities, and only asks the other three. The positive reply establishing that there is progressive disappearance of the femoral capital epiphyses lends even more weight in favour of the hypothesis of Morquio disease.

Morquio disease is suggested as follows
SMD 100%
pseudoachondroplasia 50%
Kniest disease 100%
metatropic dysplasia 100%
Refuted hypotheses
SMD
pseudoachondroplasia
Concluding Morquio disease as the diagnosis *Reason: suggested categorically*

Since Morquio disease is suggested categorically by nearly all its closest active competitors, the system correctly concludes it as the most likely diagnosis and proceeds to summarise its investigation.

You have reported the following findings
ulnae short
carpal bones small
epiphyses irregular

metacarpals proximal pointing

metacarpals short

iliac crests irregular

iliac wings wide

iliac bases hypoplastic

coxa valga

femoral necks wide

femoral capital epiphyses irregular at yrs 9

femoral capital epiphyses flat at yrs 9

femoral capital epiphyses progressive disappearance

femoral capital epiphyses abnormal from 2 yrs to 9 yrs

sacral angulation

lumbar lordosis prominent

odontoid peg absent

lumbar spine anterior beaking

vertebral end plates irregular

platyspondyly entire spine

short trunk dwarfism

male

metaphyses irregular

metaphyses sloping

The system has eliminated redundant observations from the case findings.

You have refuted the following findings

epiphyses almost normal

acetabula horizontal

iliac crests lace-like

vertebral end plates central notches

lethal

The most likely explanation of these is Morquio disease which accounts for the following findings

carpal bones small

epiphyses irregular

femoral capital epiphyses abnormal from 2 yrs to 9 yrs
femoral capital epiphyses flat at 9 yrs
femoral capital epiphyses progressive disappearance
iliac wings wide
coxa valga
lumbar spine anterior beaking
odontoid peg absent
platyspondyly entire spine
short trunk dwarfism
male
metaphyses irregular
metaphyses sloping
iliac bases hypoplastic
metacarpals proximal pointing
not acetabula horizontal
ie. 68%

The system then lists the unknown common and occasional expectations of Morquio disease and points out that excess keratin sulphate in the urine is a very significant expectation. Listing the unknown expectations is important since it informs the user of additional potential case findings which if established would further enhance the diagnosis of Morquio disease.
A feature could be unknown to the system because it has not asked about it –for example – vertebral bodies posterior scalloping. However most of the findings listed below as unknown are truly unknown because of the limited radiographs available. For example there are no chest or long bone radiographs.

Unknown expectations of Morquio disease
Common expectations
mastoid air cells under developed
mandibular condyles concave
pectus carinatum
chest AP diameter increased
ribs anterior widening

ribs posterior constriction

genu valgum

femoral capital epiphyses absent ossification from 10 yrs to 15 yrs

femoral capital epiphyses flat from 2 yrs to 10 yrs

keratin sulphate in urine excess

dentition poor

deafness progressive

AR

IQ normal

corneal opacities

ligamentous laxity

thoraco-lumbar kyphosis from 4 yrs

vertebral bodies posterior scalloping

cervical spine instability

ie 51% of the common expectations

Occasional expectations

heart enlarged

clavicles broad

diaphyseal slight widening

scoliosis

Of the unknown expectations the following are very significant

keratin sulphate in urine excess

The system finally displays all active hypotheses in descending order of explanatory power and descending order of expectations coverage (how well their entire set of expectations are satisfied, where positive occasional expectations count in favour of the diagnosis but refuted occasional expectations are ignored and do not count against the diagnosis).

Summary of consultation

Dysplasia	*Explanatory Power*	*Expectations Coverage*
Morquio disease	*68*	*44*
pseudoachondroplasia	*34*	*41*
Kniest disease	*34*	*35*
metatropic dysplasia	*25*	*21*
SMD	*24*	*31*
diastrophic dysplasia	*17*	*18*
SEDC	*12*	*12*

Morquio has the highest explanatory power. The hard abnormalities identified by the system are: metaphyses irregular and sloping, epiphyses irregular, metacarpals short and proximal pointing, femoral necks wide, platyspondyly entire spine, short trunk dwarfism, vertebral end plates irregular, iliac wings wide and carpal bones small. The hard abnormalities coverage is shown below

Dysplasia	*Hard Abnormalities Coverage*
Morquio disease	*73*
pseudoachondroplasia	*50*
Kniest disease	*45*
SMD	*45*
metatropic dysplasia	*45*
diastrophic dysplasia	*32*
SEDC	*14*

Again Morquio disease has the highest score with the highest coverage of hard abnormalities.

Appendix C: Sample diagnostic consultation using REAMS

The same series of radiographs will be used as in the sample consultation with the expert system – Appendix B (page 95).

As in the consultation with the expert system, the user has identified the following abnormalities on the radiographs:-

platyspondyly entire spine

vertebral end plates irregular

lumbar spine anterior beaking

odontoid peg absent

lumbar lordosis prominent

sacral angulation

femoral capital epiphyses flat irregular (at yrs 9)

femoral necks wide

coxa valga

iliac wings wide

iliac crests irregular

2nd – 5th metacarpals proximal pointing

epiphyses irregular

carpal bones small irregular

ulnae short

The initial consultation with REAMS is to conduct a search on the finding *platyspondyly.*

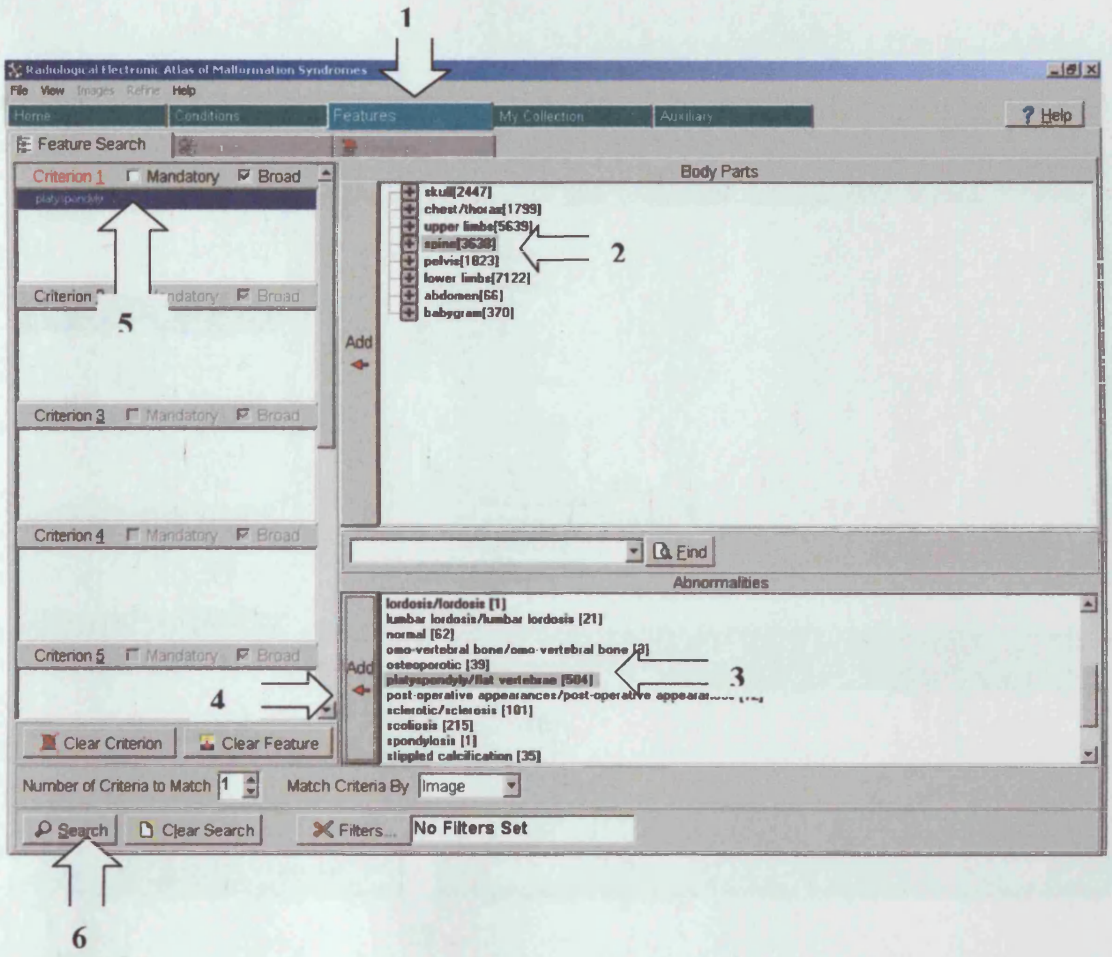
1. Open the **features** window.
2. From the **body parts** box select **spine.**
3. From the **abnormalities** box select **platyspondyly**
4. Select the **add** button adjacent to the **abnormalities** box
5. *Platyspondyly* is entered into the **criterion 1** box
6. A **search** is conducted on the single criterion. This is classified as a **broad** search, which means that all the findings which include platyspondyly are used. The finding of platyspondyly may be included as a feature of any part of the spine – cervical, thoracic, lumbar, thoraco-lumbar – and these are all used
7. The thumbnail images, all with a finding of platyspondyly, are displayed. These may be viewed individually by double clicking on each or by

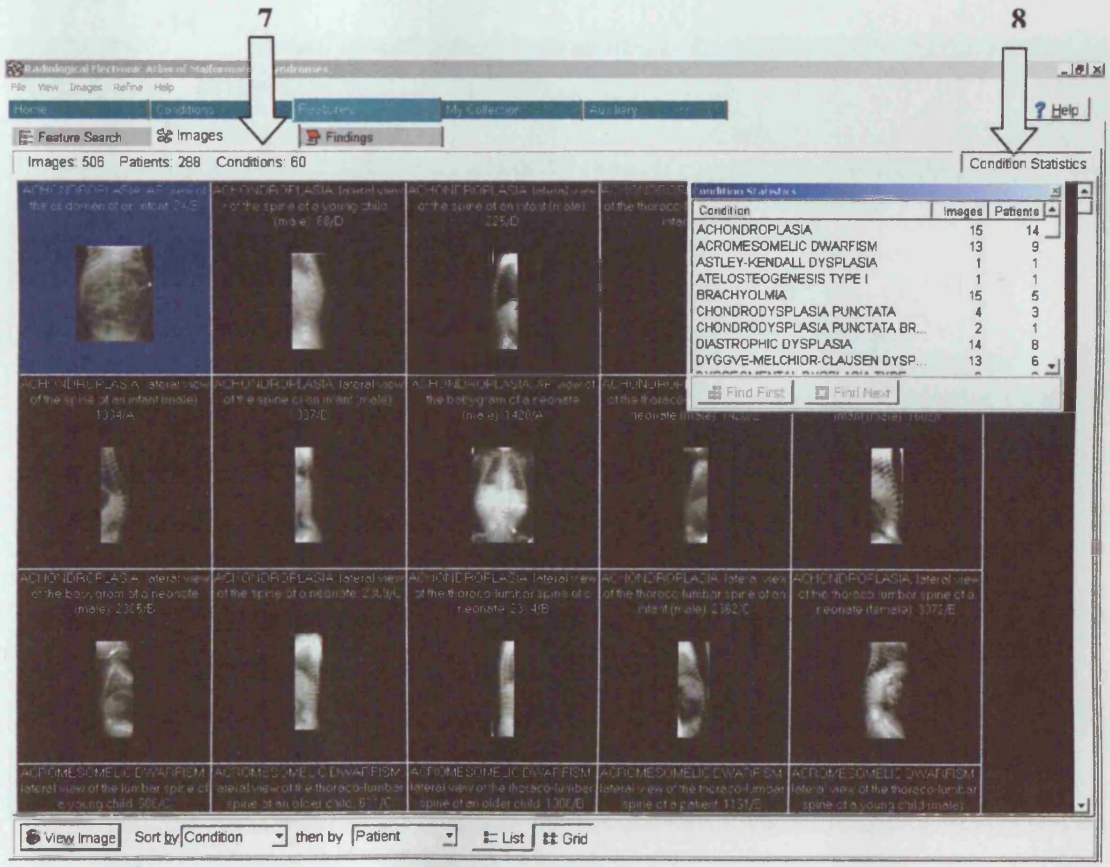
selecting **view image**. The line at the head of the images displays the number of images (506), the number of patients (288) and the number of conditions with platyspondyly (60). This represents too great a number of conditions to attempt a match with the image of the spine under investigation.

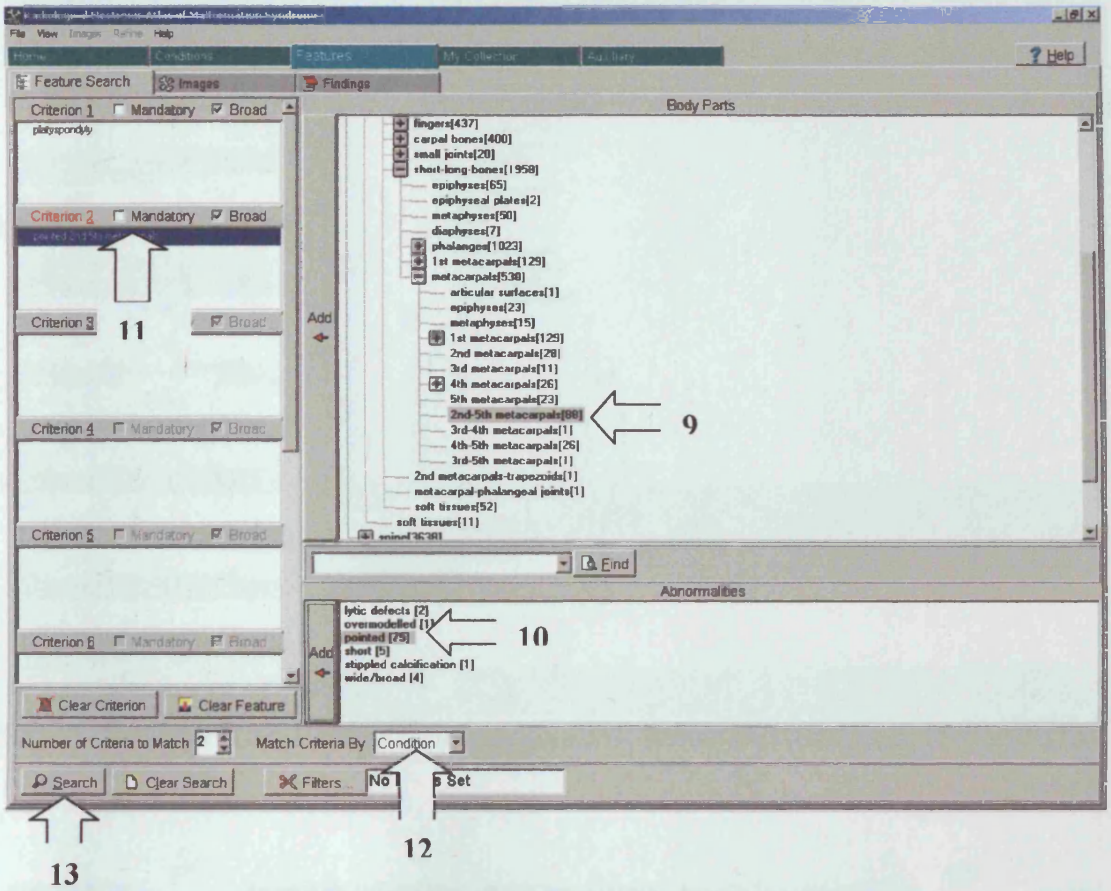
8. Selecting **condition statistics** displays in a drop-down box, the names of the dysplasias with platyspondyly and the numbers of patients and images for each diagnosis. In the **feature search** page, selecting **filters** means that the search can be refined. For example filters can be set so that the search is conducted within a given age range. In this case the age range would be between **young child** and **older child** (the available films being between the ages of 2 years and 10 years). A search with these filters set results in 33 conditions being selected with 215 images from 136 patients. This is still too large a number of diagnoses for consideration. By setting these filters, all conditions with platyspondyly illustrated only on images of the fetus, neonate or infant, are excluded. This includes all perinatally lethal dysplasias. The conditions statistics list now no longer includes the lethal conditions such as Astley-Kendall dysplasia, atelosteogenesis type 1 or fibrochondrogenesis. Returning to the **feature search** additional findings can be entered into the **criterion** boxes to refine the diagnostic possibilities.
9. Enter the finding *2^{nd} – 5^{th} metacarpals proximal pointing.* In the **body parts** box open the tree structure by clicking on **hands** then **short long bones** then **metacarpals** and finally select **2^{nd}-5^{th} metacarpals.**
10. From the **abnormalities** box select the finding closest to the observed finding – **pointed.**
11. **Add** to **criterion** box 2. The finding consisting of the combined anatomical part and the selected abnormality are displayed in the box.
12. Select **condition** rather then **images**
13. Conduct a search using the two criteria.
14. 11 dysplasias have been identified with these two criteria. There are 189 images for 99 patients. This is still not sufficiently refined to search through all the images for a visual match.
15. *Absent odontoid peg* is the third finding to be entered. An alternative to opening up the tree structure manually, is to type **odontoid peg** into the **find** box. **Go to** will automatically highlight the anatomical part in the opened

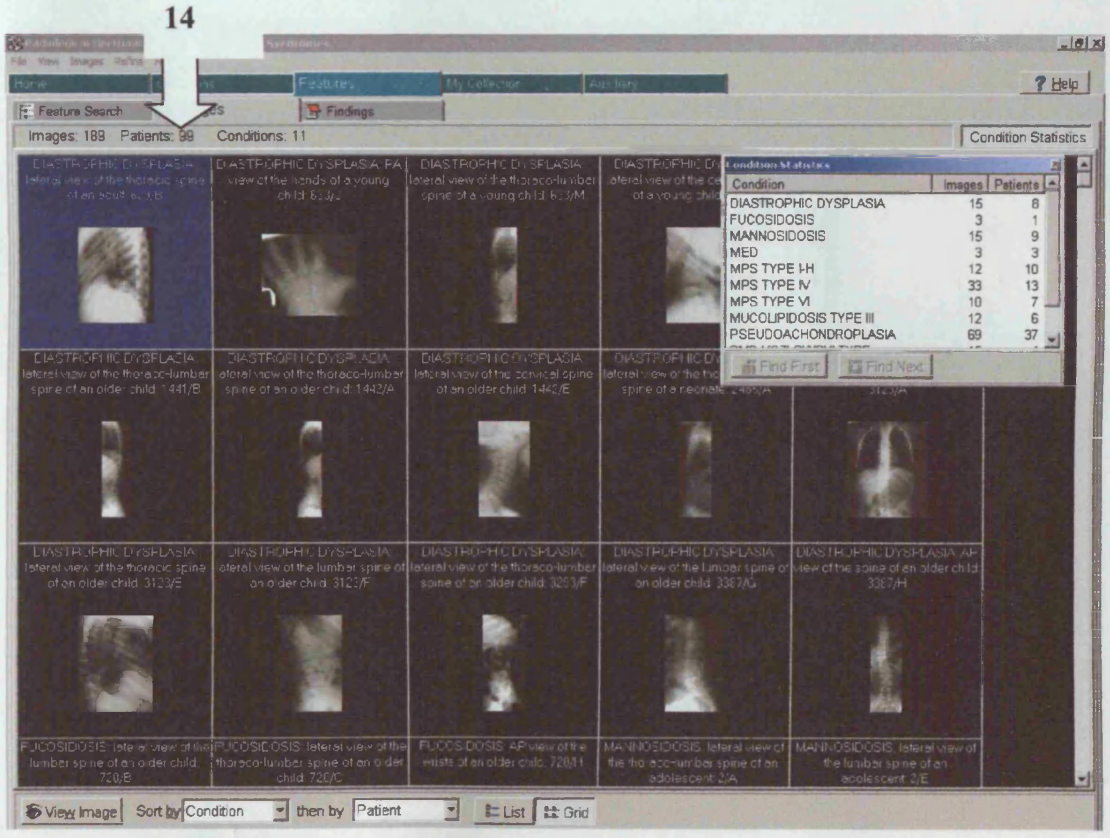
tree structure and the list of associated descriptions in the **abnormalities** box.

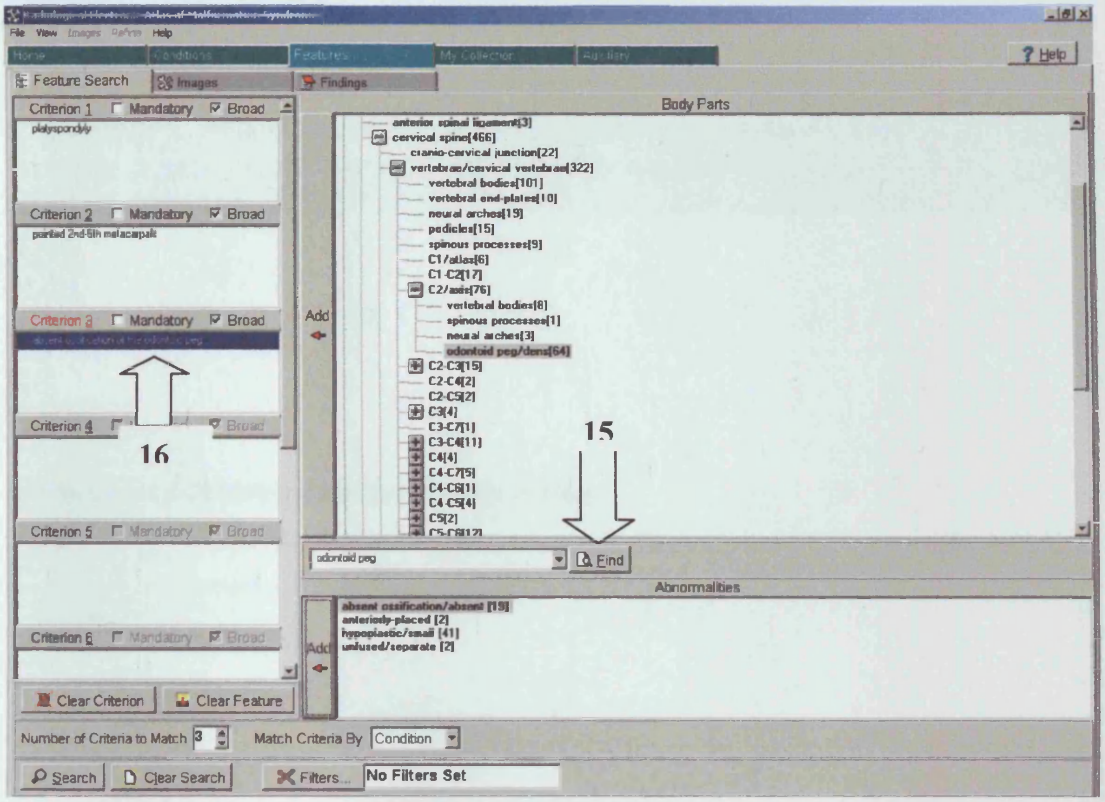
16. **Absent ossification of the odontoid peg** is added to the third **criterion** box.
17. Conducting a search demonstrates that there are now only two dysplasias with these three findings, with 49 images of 23 patients.
18. The **condition statistics** show that the two conditions are MPS type I-H and MPS type IV. The thumbnail images include views of the spine, cervical spine and hands.
19. The order or grouping of the images can be rearranged by different parameters **patient, condition, film, age, view,** with two sort parameters – **sort by** and **then by.** The images can then be viewed for a match.
20. The user would, at this point in the consultation, return to **conditions** opening up the list of conditions.
21. By selecting the tentative diagnostic possibility, and selecting **images** the full range of images associated with the condition can be viewed and compared with the other images of the limited skeletal survey for a match. A correct diagnosis of Morquio disease would then be made. In this consultation only three findings were entered to conduct a search, although many more abnormalities had been identified. In practice this is how the system should be used so that the full details of a few differential diagnoses can be compared.
22. By selecting **findings**, the list of findings associated with images may be viewed. These are presented in three possible formats, **short, standard, full** referring to the reporting form. In the full reporting form qualifiers are included (mildly, proximally, distally) but are excluded in the short reporting format. Some of the findings become merged in the short reporting form making the list of findings also shorter. When conducting a feature search, short findings are displayed. For example the user identified one finding as *2^{nd} – 5^{th} metacarpals proximal pointing. Proximal* is not included in the feature list and the closest finding is *2^{nd} – 5^{th} metacarpals pointed.*

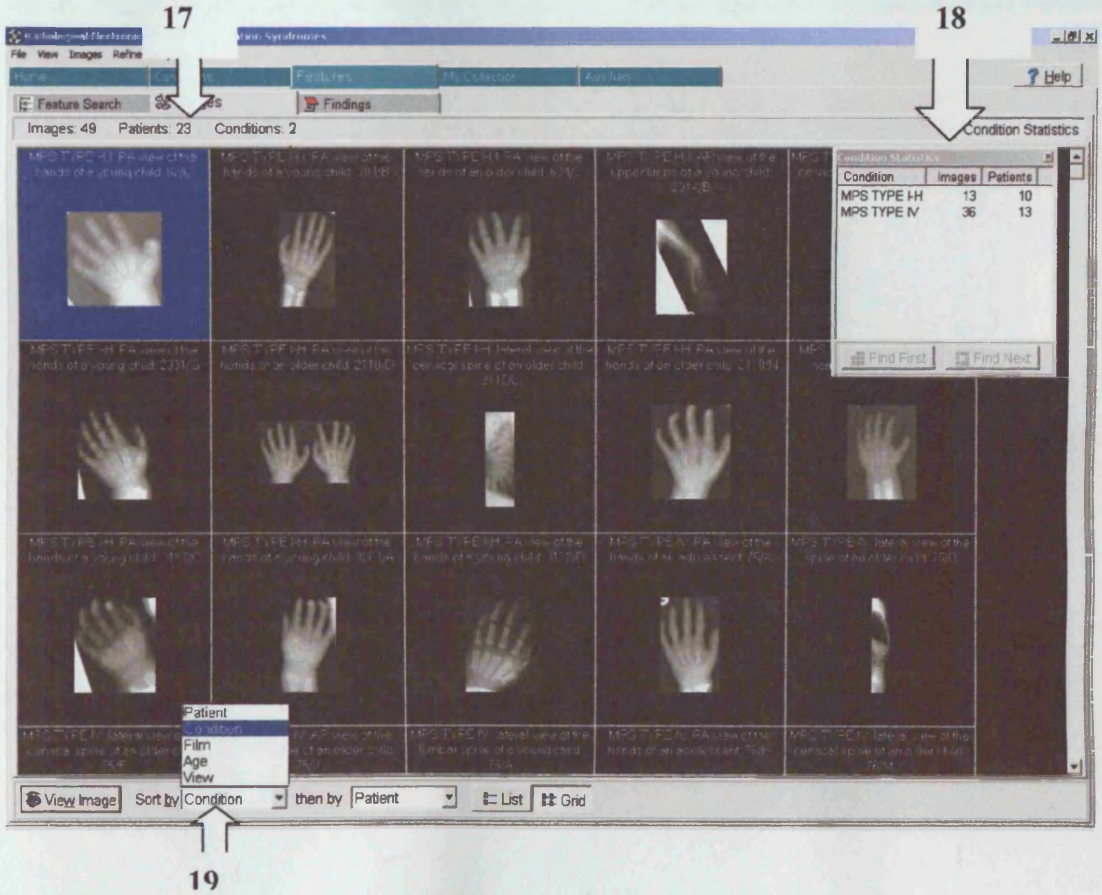

1
Radiological Electronic Atlas of Malformation Syndromes
File View Images Refine Help
Home
Conditions
Features
My Collection
Auxiliary
Help
Feature Search
Criterion 1
Mandatory
Broad
Criterion 3
Criterion 4
Criterion 5
Clear Criterion
Clear Feature
Body Parts
skull[2447]
chest/thorax[1799]
upper limbs[5639]
spine[3628]
pelvis[1823]
lower limbs[7122]
abdomen[66]
babygram[370]
Add
Find
Abnormalities
lordosis/lordosis [1]
lumbar lordosis/lumbar lordosis [21]
normal [62]
osteoporotic [39]
platyspondyly/flat vertebrae [504]
sclerotic/sclerosis [101]
scoliosis [215]
spondylosis [1]
stippled calcification [35]
2
3
4
5
6
Number of Criteria to Match
Match Criteria By
Image
Search
Clear Search
Filters...
No Filters Set

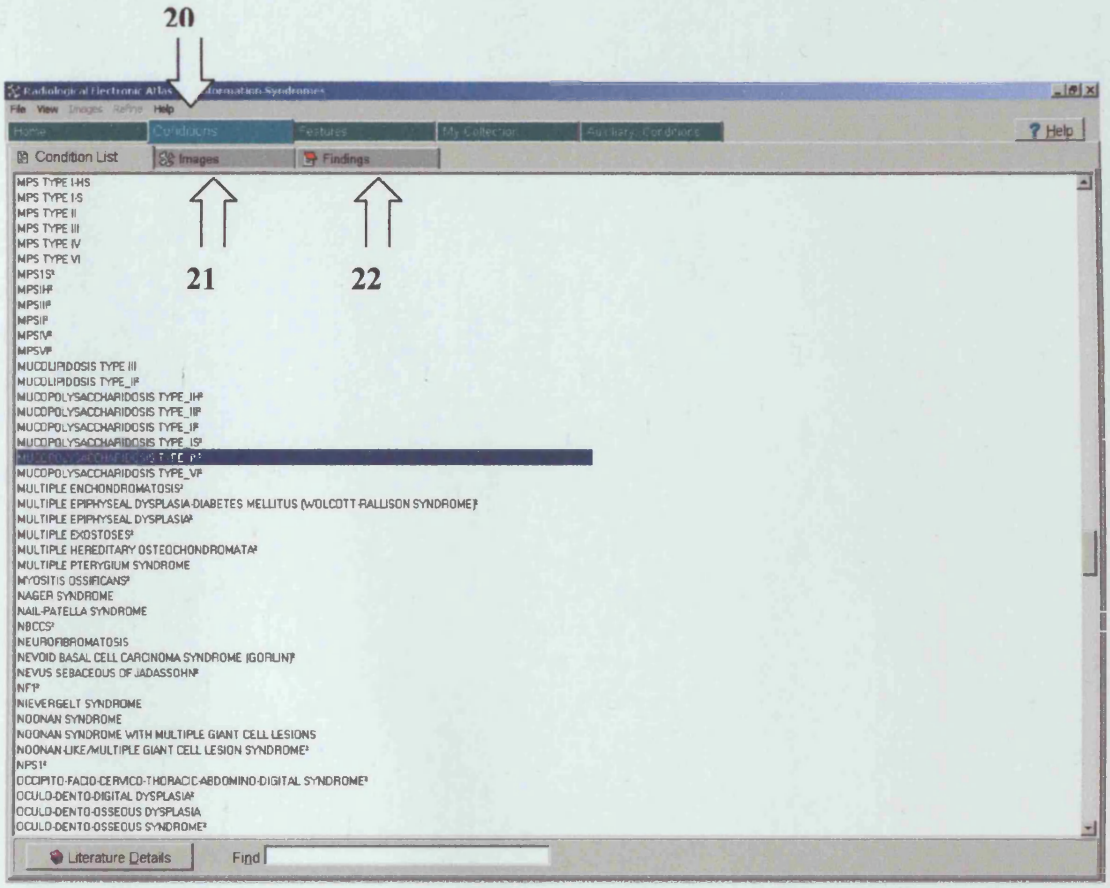

7
8
Feature Search
Images
Findings
Images: 506 Patients: 288 Conditions: 60
Help
Condition Statistics
Condition
Images
Patients
ACHONDROPLASIA 15 14
ACROMESOMELIC DWARFISM 13 9
ASTLEY-KENDALL DYSPLASIA 1 1
ATELOSTEOGENESIS TYPE I 1 1
BRACHYOLMIA 15 5
CHONDRODYSPLASIA PUNCTATA 4 3
CHONDRODYSPLASIA PUNCTATA BR... 2 1
DIASTROPHIC DYSPLASIA 14 8
DYGGVE-MELCHIOR-CLAUSEN DYSP... 13 6
Find First
Find Next
View Image
Sort by Condition
then by Patient
List
Grid


Feature Search
Images
Findings
Criterion 1
Mandatory
Broad
platyspondyly
Criterion 2
Criterion 3
Criterion 4
Criterion 5
Criterion 6
11
Body Parts
fingers[437]
carpal bones[400]
small joints[20]
short-long-bones[1958]
epiphyses[65]
epiphyseal plates[2]
metaphyses[50]
diaphyses[7]
phalanges[1023]
1st metacarpals[129]
metacarpals[538]
articular surfaces[1]
epiphyses[23]
metaphyses[15]
1st metacarpals[129]
2nd metacarpals[28]
3rd metacarpals[11]
4th metacarpals[26]
5th metacarpals[23]
2nd-5th metacarpals[98]
3rd-4th metacarpals[1]
4th-5th metacarpals[26]
3rd-5th metacarpals[1]
2nd metacarpals-trapezoids[1]
metacarpal-phalangeal joints[1]
soft tissues[52]
soft tissues[11]
9
Add
Find
Abnormalities
lytic defects [2]
overmodelled [1]
pointed [25]
short [5]
stippled calcification [1]
wide/broad [4]
10
Clear Criterion
Clear Feature
Number of Criteria to Match
Match Criteria By
Condition
Search
Clear Search
Filters
Set
13
12


14
Feature Search
Findings
Images: 189
Patients: 99
Conditions: 11
Condition Statistics
Condition
Images
Patients
DIASTROPHIC DYSPLASIA 15 8
FUCOSIDOSIS 3 1
MANNOSIDOSIS 15 9
MED 3 3
MPS TYPE I-H 12 10
MPS TYPE IV 33 13
MPS TYPE VI 10 7
MUCOLIPIDOSIS TYPE III 12 6
PSEUDOACHONDROPLASIA 69 37
Find First
Find Next
View Image
Sort by
Condition
then by
Patient
List
Grid


Feature Search
Images
Findings
Help
Criterion 1 Mandatory Broad
platyspondyly
Criterion 2 Mandatory Broad
Criterion 3 Mandatory Broad
Criterion 4 Mandatory Broad
Criterion 5 Mandatory Broad
Criterion 6 Mandatory Broad
Clear Criterion
Clear Feature
Body Parts
anterior spinal ligament[3]
cervical spine[466]
cranio-cervical junction[22]
vertebrae/cervical vertebrae[322]
vertebral bodies[101]
vertebral end-plates[10]
neural arches[19]
pedicles[15]
spinous processes[9]
C1/atlas[6]
C1-C2[17]
C2/axis[76]
vertebral bodies[8]
spinous processes[1]
neural arches[3]
odontoid peg/dens[64]
C2-C3[15]
C2-C4[2]
C2-C5[2]
C3[4]
C3-C7[1]
C3-C4[11]
C4[4]
C4-C7[5]
C4-C6[1]
C4-C5[4]
C5[2]
Add
odontoid peg
Find
Abnormalities
absent ossification/absent [19]
anteriorly-placed [2]
hypoplastic/small [41]
unfused/separate [2]
Number of Criteria to Match 3
Match Criteria By Condition
Search
Clear Search
Filters...
No Filters Set
16
15


17
18
Help
Feature Search
Findings
Images: 49 Patients: 23 Conditions: 2
Condition Statistics
Condition Images Patients
MPS TYPE I-H 13 10
MPS TYPE IV 36 13
Find First
Find Next
Patient
Condition
Film
Age
View
View Image
Sort by Condition
then by Patient
List
Grid
19


20
Radiological Electronic Atlas
File View Images Refine Help
Home
Conditions
Features
My Collection
Auxiliary Conditions
Help
Condition List
Images
Findings
21
22
MPS TYPE I-HS
MPS TYPE I-S
MPS TYPE II
MPS TYPE III
MPS TYPE IV
MPS TYPE VI
MPS1S
MPSIH
MPSII
MPSIS
MPSIV
MPSVI
MUCOLIPIDOSIS TYPE III
MUCOLIPIDOSIS TYPE_II
MUCOPOLYSACCHARIDOSIS TYPE_IH
MUCOPOLYSACCHARIDOSIS TYPE_II
MUCOPOLYSACCHARIDOSIS TYPE_IS
MUCOPOLYSACCHARIDOSIS TYPE_VI
MULTIPLE ENCHONDROMATOSIS
MULTIPLE EPIPHYSEAL DYSPLASIA-DIABETES MELLITUS (WOLCOTT-RALLISON SYNDROME)
MULTIPLE EPIPHYSEAL DYSPLASIA
MULTIPLE EXOSTOSES
MULTIPLE HEREDITARY OSTEOCHONDROMATA
MULTIPLE PTERYGIUM SYNDROME
MYOSITIS OSSIFICANS
NAGER SYNDROME
NAIL-PATELLA SYNDROME
NBCCS
NEUROFIBROMATOSIS
NEVOID BASAL CELL CARCINOMA SYNDROME (GORLIN)
NEVUS SEBACEOUS OF JADASSOHN
NF1
NIEVERGELT SYNDROME
NOONAN SYNDROME
NOONAN SYNDROME WITH MULTIPLE GIANT CELL LESIONS
NOONAN-LIKE/MULTIPLE GIANT CELL LESION SYNDROME
NPS1
OCCIPITO-FACIO-CERVICO-THORACIC-ABDOMINO-DIGITAL SYNDROME
OCULO-DENTO-DIGITAL DYSPLASIA
OCULO-DENTO-OSSEOUS DYSPLASIA
OCULO-DENTO-OSSEOUS SYNDROME
Literature Details
Find

A Model-Based Diagnostic Expert System for Skeletal Dysplasias

E.T. Keravnou, J. Washbrook
Department of Computer Science, University College London
Gower Street, London WC1E 6BT, UK

R.M. Dawood, C.M. Hall, D. Shaw
Department of Radiology, The Hospital for Sick Children
Great Ormond Street, London WCN 3JH, UK

Abstract

A prototypical model-based diagnostic expert system for skeletal dysplasias is discussed in the context of the competent expert systems methodology and an advanced generic architecture for second generation diagnostic systems.

key words: model-based diagnostic system, second generation diagnostic architecture, cooperating expert systems, competent expert systems methodology.

A Model-Based Diagnostic Expert System for Skeletal Dysplasias

Abstract

A prototypical model-based diagnostic expert system for skeletal dysplasias is discussed in the context of the competent expert systems methodology and an advanced generic architecture for second generation diagnostic systems.

key words: model-based diagnostic system, second generation diagnostic architecture, cooperating expert systems, competent expert systems methodology.

1. INTRODUCTION

The objectives of this paper are twofold: firstly, to present SDD (Skeletal Dysplasias Diagnostician), a prototypical model-based diagnostic expert system for skeletal dysplasias constructed through the competent expert systems methodology (Keravnou and Johnson, 1986); secondly, to use this to illustrate an advanced generic architecture for second generation diagnostic systems. First generation diagnostic systems have serious limitations (Bell, 1985; Davis, 1982; Dhar and Pople, 1987; Clancey, 1983; Keravnou and Johnson, 1986; Kidd and Cooper, 1985) which probably explains why medical expert systems have not so far been accepted by the medical community.

The Problem Domain

A skeletal dysplasia is a generalised disorder affecting the growth of bone and cartilage. The diagnosis of dysplasias from X-ray films is a skilled task. Individual dysplasias are relatively rare and consequently expertise is scarce. For the parents of an affected child, knowledge of the prognosis and best-known treatment for the dysplasia is of great value. Equally valuable is genetic counselling, to inform the parents of the chances of other offspring being similarly affected.

The Hospital for Sick Children in Great Ormond Street (GOS) is a centre for referral from all over the world. The expert system aimed to be built will model the diagnostic skills of GOS; it will use an on-line video library of X-ray images for an easier and more reliable categorisation of features and signs. The system will make the relevant expertise widely and cheaply available with immense social and financial implications. In addition to being a diagnostic aid to the radiologists who are not expert in dysplasias the system will also aid the experts at GOS in recognizing new dysplasias. This paper is only concerned with the diagnostic system, SDD.

The Competent Expert Systems Methodology

SDD was built by applying the competent expert systems methodology (Keravnou and Johnson, 1986). Central to this methodology is the process of knowledge elicitation. Although the terms knowledge elicitation and knowledge acquisition are often used

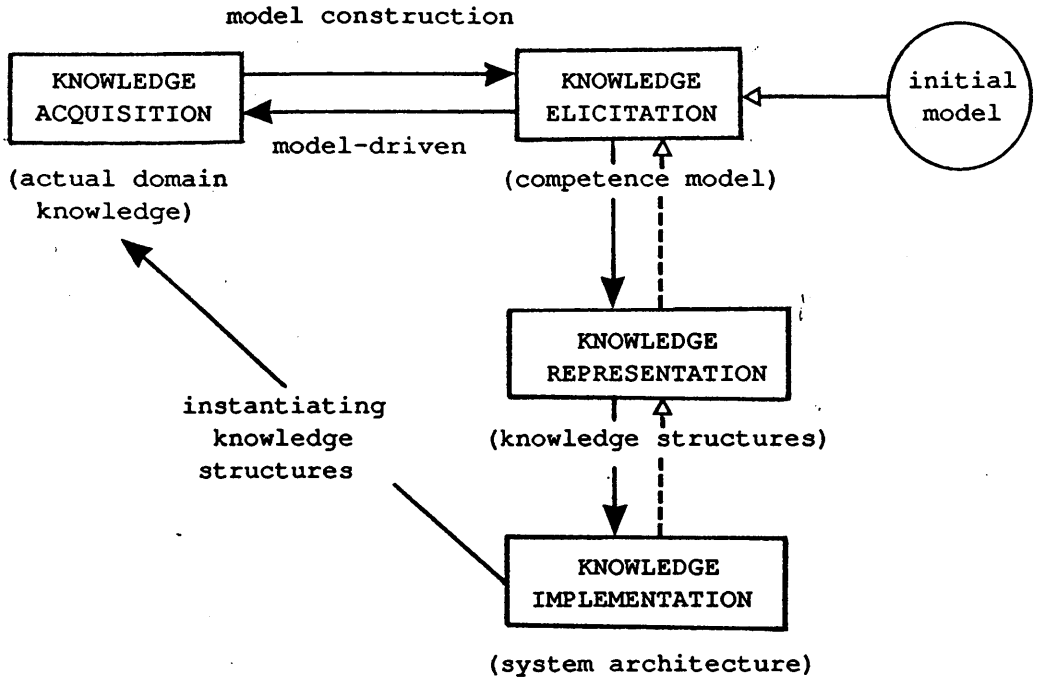


Figure 1 Processes in Competent Expert Systems Methodology

interchangeably in the literature, the methodology perceives a clear distinction between them. The two processes are closely related which accounts for their being confused and makes the distinction all the more necessary.

Knowledge elicitation is the process of formulating a model, the *competence model*, for the particular expertise. The competence model consists of:

- A model of the factual knowledge of the domain. The model specifies the entity classes, their properties and their interrelationships. Thus the model specifies the conceptual organisation of the factual knowledge.
- Models for the domain reasoning processes.
- Abstract expressions of reasoning strategies.

Knowledge acquisition is the process of obtaining actual domain knowledge (both factual knowledge and reasoning knowledge).

Knowledge elicitation and knowledge acquisition are intimately related (see figure 1). The existence of an accurate competence model for a particular domain provides the basis for a focussed acquisition of domain knowledge; in a sense any knowledge that does not conform to the model is not relevant. This is referred to as *model-driven* knowledge acquisition.

However, possessing such a model of competence prior to embarking on the task of acquiring the domain knowledge is rather unlikely. Usually the formulation of the model and the acquisition of knowledge proceed incrementally and in parallel, the one process reinforcing the other: from some initial domain knowledge an initial competence model, which is probably inaccurate and incomplete, is formulated. This model can form the basis for acquiring more knowledge, thus testing its accuracy (coverage and resolution). The more knowledge that is acquired which conforms to the model the stronger the belief in the model becomes. On the other hand when a newly acquired piece of knowledge does not conform to the model, the model may need to be extended or modified, always safeguarding against the possibility of irrelevant knowledge.

It is, therefore, important that the knowledge engineer formulates a model of competence for the particular domain very early on, even if the model is subsequently radically revised.

Advanced Diagnostic Architecture

In a diagnostic task the important concepts are *findings* (symptoms, signs, historical data, laboratory data, radiological signs) and *hypotheses* (explanations of abnormal findings). The proposed advanced architecture (Keravnou and Johnson, 1988) separates these two bodies of factual knowledge. A *findings reasoner* operates on the findings knowledge in order to make intelligent inferences on the available case-specific information. Such inferencing could be of a common-sense nature or it could be based on specialist knowledge. A *diagnostician* uses the hypotheses knowledge to generate and refine case-specific hypotheses. The *diagnostic picture* is the global data structure holding the results of both the findings reasoner and the diagnostician.

The prototype diagnostic system SDD will be discussed in the framework of the competent expert systems methodology by presenting the elicited competence model and the knowledge structures designed for it. The architecture of SDD will be presented in the background of the advanced diagnostic architecture outlined above.

2. OVERVIEW OF THE SDD COMPETENCE MODEL

In this section the terms *features, findings, subjects* and *attributes* are defined, followed by a brief description of the SDD Competence Model.

Features and Findings

A dysplasia is described by its *features*. Some examples of features are Flared metaphyses, Short limbs from birth, Severe myopia, Platyspondyly, and Coxa vara. Each of these features has a *subject*, each of which may have one or more *attributes*. A feature is a concise description, and its conciseness relies upon the assumption of a general background knowledge. For example, the feature Flared metaphyses (subject: metaphyses; attribute: shape, value: flared) says that some or all of the long bones have flared metaphyses. The feature is to be interpreted in the context of growth regions in long bones. For doctors this interpretation comes from their general medical knowledge and reference material. For a computer system the medical knowledge needs to be explicitly represented, and this is achieved through the feature model. The system is also able to use this knowledge to provide explanations of features to users. Part of the knowledge is taxonomic, for example the vertebrae are part of the spine, which is part of the skeleton. Thus if the finding Spine normal is reported it can be deduced (by the findings reasoner) that platyspondyly, a flattening of some or all of the vertebrae, is absent.

When a radiologist is presented with a case usually some features are immediately apparent to the trained eye. These initial *findings*, and other findings revealed by further examination and investigation, form the basis upon which a diagnosis is made. Thus a dysplasia is described by its features, and a case by its findings – although as the findings which are relevant to a diagnosis will become features of the case, the terms are loosely used interchangeably.

The SDD Competence Model

The Competence Model for SDD comprises two major components, the Domain Knowledge Model and the Diagnostic Model (figure 2). The Domain Knowledge Model in turn consists of two models, a Dysplasia Model and a Feature Model. The Dysplasia Model is a description of a dysplasia in terms of its features and its relations with other dysplasias. The Feature Model is a description of features and is used by the Findings Reasoner.

Competence Model

- Domain Knowledge Model
  - — Dysplasia Model
  - — Feature Model
- Diagnostic Model
  - — Diagnostic Procedure
  - — Hypothesis Status-Transition Model
  - — Findings Reasoner

Figure 2. The Competence Model

The Diagnostic Model has a Diagnostic Procedure which matches findings against dysplasias, and generates and evaluates hypotheses about dysplasias. It uses the Domain Knowledge Model and the Findings Reasoner. Whereas the knowledge in the Diagnostic Procedure is domain-specific, the Findings Reasoner contains general medical knowledge and is able to make the kind of common-sense deductions such as the one about platyspondyly above.

At any given time a number of hypotheses may be entertained. The Hypothesis Status-Transition Model is used to record the histories of the various hypotheses considered during a consultation, which provide the bases for explaining diagnoses.

3. DOMAIN KNOWLEDGE MODEL

Dysplasia Model

Each dysplasia (and group of dysplasias) is characterised by 5 sets of features, which in decreasing order of diagnostic power are:

- *Typical:* Possibly rare features (eg 1% of cases) but conclusive even when occurring relatively infrequently; the absence of typical features from a case must not count against the possibility of the dysplasia. (A typical feature is diagnostic in the context of some other evidence that there is an abnormality; for radiological cases this will necessarily be the case, otherwise no X-ray would have been taken.)
- *Sufficient:* The meaning of a sufficient set is that if the particular set of features is observed, the associated dysplasia can be established as a *working hypothesis*. (It does not mean sufficient to confirm as a diagnosis.) There could be a number of sufficient sets for a dysplasia; these would be expected to intersect and their intersection set would include those features which are absolutely necessary for the dysplasia to occur. Sufficient sets are probably subsets of common (see below).
- *Triggers:* The purpose of a trigger is to direct attention economically towards *possible* diagnoses. Again there could be a number of triggers for a dysplasia. Triggers are of two types:

— *Primary triggers:* Features that catch the radiologist's attention, which could include clinical observations, discriminatory information like age or sex, and radiological features.

— *Secondary triggers:* When a possibility or a working hypothesis is being explored, new observations are expected to refer to features associated with the particular dysplasia. If "unexpected" responses (not supporting the pursued possibility or hypothesis) are given these may suggest another possible dysplasia. Such responses are secondary triggers.

- *Common:* Features which have been observed to occur in the majority of cases. Given that the patient suffers from this dysplasia one would expect to observe the common features. Their absence would need to be taken into account.
- *Other:* Observed co-incidental abnormalities currently having no diagnostic significance (effectively allowing them to be ignored in the final diagnosis but also preventing their presence from counting against the hypothesis of the dysplasia). Through diagnostic experience, "other" features may be upgraded to common or typical, thus allowing for an evolutionary system.

Feature Model

In addition to the attributes and taxonomy of finding subjects mentioned in the overview in section 2 a number of other aspects of features are modelled, and these are discussed below.

Finding Subject Taxonomy

The subjects of findings (*finding subjects*) are related in taxonomies (see figure 3). For example, the vertebrae, spine and skeleton are in a bone taxonomy. A taxonomy of finding subjects not only allows the generalisation or restriction of findings but also the sharing of common characteristics (attributes, values etc) and the modelling of exceptions.

Attributes of Subjects

Most attributes, such as sex, age, location, and size, are single-valued although that value may take a number of forms, eg age could be expressed qualitatively (stillborn, baby, infant, child, adult) or numerically (0-6 months, 2 years, over 5 years). Multi-valued attributes are also possible, eg the shape of a particular bone could be reported as "long and thin".

Consideration also has to be given to the relation between findings, for example the negative finding Femoral-head absent makes it pointless to ask about the location of the femoral head.

Multiple X-ray Views

X-ray films give a particular view of a bone, eg lateral view or frontal view. If different X-ray views of the same bone are discussed using different attributes and/or

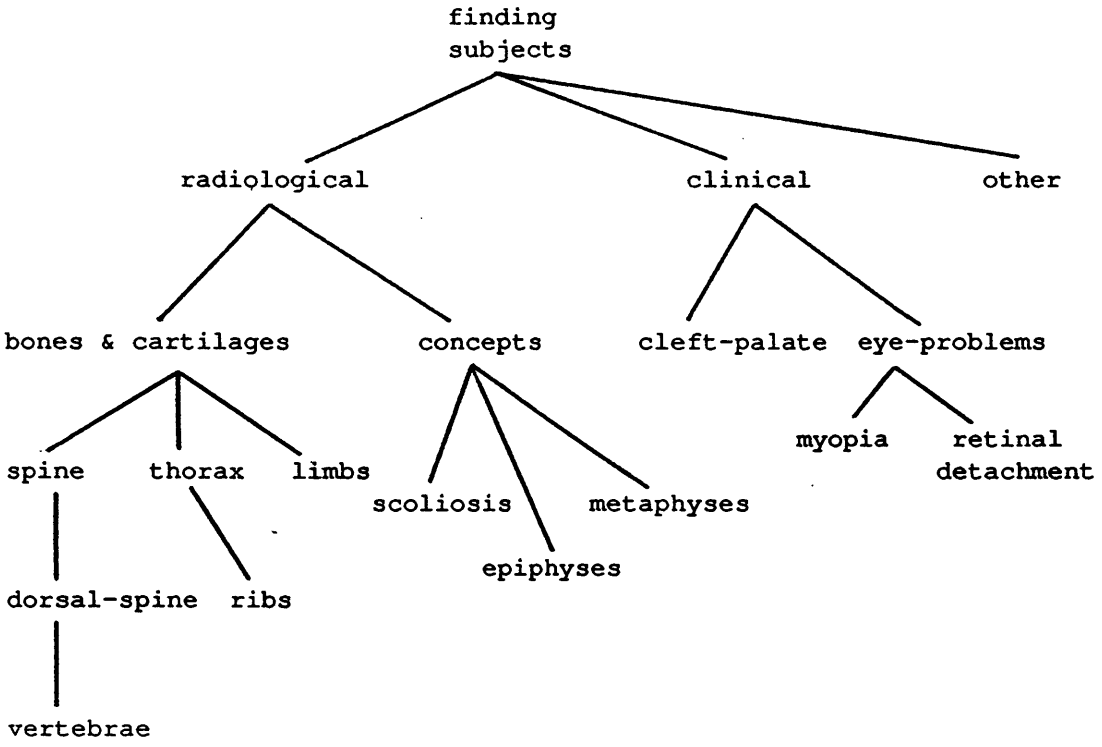


Figure 3 Taxonomy of Finding Subjects

values then the model for that bone must make explicit these different view perspectives. In addition it must specify correspondences, if any, between these views.

Spatial Relations

Bones are spatially related. In skeletal dysplasias normal spatial relations become distorted. *Location* is the attribute describing the position of the bone in relation to other bones or the description of the displacement from normality.

Another related issue is the case of parts of the skeleton which are made up of a number of bones of a similar kind, eg the dorsal spine has 12 vertebrae and the thorax contains 12 pairs of ribs. Radiological features describe abnormalities with reference to the component bones (eg platyspondyly) and qualify them by giving the portion (in qualitative terms) of the part affected (eg affects entire spine). Attribute *affects* describes qualitatively the extend of the abnormality on the part of the skeleton for which values could be specific to the particular part of the skeleton, eg localities on the thorax could be described as left and right, whilst localities within the dorsal spine could be described as upper, mid and lower.

Temporal Relations

In any medical domain findings are qualified by temporal aspects. In the simplest situation every finding holds *currently*. In the domain of skeletal dysplasias modelling time is central to the problem since dysplasia manifestations vary significantly with age. Salient features can disappear with age (which makes an accurate early diagnosis all the more necessary) while others can only be detected after a certain age (eg progressive kyphoscoliosis cannot always be detected from birth). The current model of time is rather rudimentary. To facilitate extensions, all reasoning about time will be the task of a separate module.

In the example "Short stature from birth" the temporal information is expressed relative to the patient age. Such temporal specifications are referred to as *absolute temporal aspects* and temporal relations between findings as *relative temporal aspects* (cf Allen's temporal logic (Allen, 1983 & 1984)). Currently only absolute temporal aspects are modelled, by expressing the interval in the lifetime of a patient during which a feature holds. Time intervals are expressed by their start and end points or simply by their start points (see above example) indicating an on-going situation. Since patient age will be grossly expressed as say *stillborn, infant, x months, x years,* it would be more natural to give the duration of the occurrence of the feature instead of the end point of the time interval. In addition the pattern of occurrence eg continously, intermittently etc may be expressed, although this would be more appropriate for clinical data.

Absence of explicit temporal information would normally default to *currently*. With stillborn cases temporal reasoning reverts to the default situation.

Correspondence between Radiological and Clinical data

Short forearm is a clinical finding with possible radiological correlates Short radius and ulna. However both essentially describe the same thing. Such correspondences between clinical and radiological data need to be modelled.

Radiological Concepts

The radiological subjects consist of specific bones (and cartilages) and more abstract concepts. Presently a concept is meant to be anything that describes or defines some aspect of bones and applies not just to a single bone. This is best explained by some examples. Consider the features, Knee epiphyses not present at birth, Wide metaphyses, Flat epiphyses. Each long bone has an epiphysis and a metaphysis. The first finding above makes the context of epiphyses explicit, namely the knee. The other two findings potentially refer to all long bones. Representing epiphyses and metaphyses as parts of individual bones would not allow this level of abstraction.

Knowledge Structures

In this section the knowledge structures for dysplasias and finding subjects are overviewed. In both cases the primary representation scheme is frames; pseudo-Lisp notation is used for expressing these structures.

```
(<dysplasia-name>
  (long-name <text>)
  (dysplasia-group <dysplasia-name>)
  (typical <feature 1> ... <feature n>)
  (triggers <trigger-name 1> ... <trigger-name n>)
  (sufficient
        (<feature 1,1> ... <feature 1,n>)
                .........
        (<feature m,1> ... <feature m,n>))
  (common <feature 1> ... <feature n>)
  (other <feature 1> ... <feature n>)
  (differential-diagnosis
        (<dysplasia-name> <feature 1> ... <feature n>)
                  .........
        (<dysplasia-name> <feature 1> ... <feature n>))
  (refinements <dysplasia-name 1> ... <dysplasia-name n>)
  (refinement-suggestions
        (<refinement>
             (<feature 1> ... <feature n>) ...
             (<feature 1> ... <feature n>))
        (<refinement> ... ) ....))
```

Figure 4 Dysplasia Frame Format

Dysplasia Structure

Referring to figure 4 slots *common*, *typical*, *sufficient* and *other* refer to the corresponding feature sets for the dysplasia. The *triggers* slot represents the primary triggers for the dysplasia and the *differential-diagnosis* slot the secondary triggers associated with the dysplasia. *Dysplasia-group* and *refinements* slots provide upwards and downwards taxonomic relations respectively. The *refinement-suggestions* slot gives sets of features for selecting a refinement. A dysplasia inherits common features from its group.

Dysplasia manifestations vary significantly with age. Initially it was decided to represent the age-independent profile of the dysplasia in one frame and age-specific profiles in subordinate frames (inheriting information from the former). Soon it became apparent that this rather complex arrangement was unnecessary; since features have explicit temporal aspects, given the patient age it is easy to screen out the features that do not apply. For example given a five month old baby one cannot talk about flared metaphyses at the age of 2.

Finding Subject Structure

The finding subject structure is given in figure 5. Slot *trigger-parts* identifies the primary triggers which involve findings of the particular subject. Slots *isa* and *part-of* represent the relevant taxonomic links with respective inverse links given in slots *type-instances* and *components* (in the latter case the number of each component type is specified).

Slot *prompting-questions* has two uses at the moment: first for making a user-volunteered finding more specific, eg if the user enters Coxa vara, the system may attempt to make this finding more specific by asking about the severity and lateracy of coxa vara; secondly for establishing correspondences between clinical and radiological findings, eg Short stature may be because the femur is displaced, and this can be asked. In either case a finding activates a procedure embodying the appropriate questioning sequence.

Slot *to-instantiate* represents the procedure for instantiating the attributes for the particular subject. Specific procedures will make use of the individual attribute descriptors, such as *if-needed* procedures (see below) and explanatory information (see next). Because skeletal dysplasias are individually rare, there is no standard, widely known, terminology. For example fibrochondrogenesis is unlikely to be known to a radiologist who is not an expert in this field. Such features are better explained through X-ray images or line diagrams. Slot *to-explain* does this by representing a procedure that displays either text (eg cystic masses is better known as cauliflower ears) or a diagram.

The attribute descriptors are included in slot *attributes* which also specifies the negative finding of the subject. Each attribute descriptor specifies the type of the attribute and its value-set, the values referring to normality, the default value, synonyms, and a procedure (*if-needed*) which may determine a value for the attribute. *Abstractions* represent mappings from quantitative to qualitative values for the

attribute (this could be represented in terms of a procedure or as a collection of rules).

Dependencies between a finding and other findings of the same or different subjects are represented in slot *implications*.

Slot *x-ray-views* is only relevant to bone subjects. It links the subject frame to frames providing X-ray perspectives for that subject.

Lastly slot *detectability* associates findings for that subject with age intervals during which they can be detected; again this is relevant to bone subjects.

```
(<finding-subject>
   (trigger-parts <trigger-name 1> ... <trigger-name n>)
   (isa <finding-subject>)
   (type-instances <finding-subject 1> ... <finding-subject n>)
   (part-of <finding-subject> <number-of>)
   (components
      (<number> <finding-subject 1>) .....
      (<number> <finding-subject n>))
   (prompting-questions
      (<finding> [<seek-to-establish-finding>] <procedure>) ... )
   (to-instantiate <procedure>)
   (to-explain
      (<finding> <procedure>) ... )
   (attributes
      (negative-finding <attribute> <value>)
      (<attribute 1>
         (type mv/sv)
         (value-set <value 1> ... <value n>)
         (normal <value/s>)
         (default <value/s>)
         (abstractions <procedure>/<rules>)
         (synonyms (<value> <synomyms>) ... )
         (if-needed <procedure>))
           .......
      (<attribute n> .....))
   (implications
      (<required-finding> <implicated-finding 1> ... <implicated-finding n>)
          .... )
   (x-ray-views
      (<view> <view perspective frame for finding subject>) ...))
   (detectability
      (<finding> <age-interval>)))
```

Figure 5 Finding Subject Frame

Paradoxically, the finding subject structure is more complicated than the dysplasia structure. The complexity is due to the open-ended nature of the reasoning task (findings reasoner) which operates on the finding subject frames. Aspects of this reasoning will be of a common-sense nature which makes it all the more difficult to delineate. In contrast, being of a specialist nature, the diagnostic reasoning is easier to

delineate and the associated knowledge structures are comparatively simple. The separation of the findings knowledge from the dysplasia knowledge is responsible for much of the simplicity of the dysplasia structure (see section 5).

Finding Format

Simple findings are expressed in the format:

```
(<subject> (<attribute> [or] <values>) ...
           (<attribute> ...)
           (time <interval>))
```

Compound findings are conjunctions or disjunctions of findings, eg the combination Well-developed acetabulum with upward femoral displacement is very indicative of SEDC whilst neither of these features on their own is. Also in SEDC pubic bones may be absent or short. This is expressed as (pubic-bones (status or absent short)), which is more elegant than the compound finding (or (pubic-bones (status absent)) (pubic-bones (status short))).

The finding format as given is suitable for an internal representation. The user should be able to enter findings in a more flexible and friendly way, possibly though a natural language interface.

Trigger Frame

All the primary triggers are kept in a single frame structure which has a slot for every trigger as:

```
((<trigger-name 1> <dysplasia-name> <findings>)
  ....
 (<trigger-name n> ...))
```

4. DIAGNOSTIC MODEL

The diagnostic model is a model of the reasoning involved in the particular diagnostic task. It therefore provides a dynamic view of the problem domain in contrast to the static view provided by the domain knowledge model.

Figure 6 is a flow diagram of the diagnostic process at a high level of abstraction. This diagram was constructed and refined in parallel with the construction of the dysplasia model. After a few sessions with the experts (both going through cases and discussing dysplasias independently of actual cases) it was possible to draft preliminary dysplasia and diagnostic models. The preliminary diagnostic model was deliberately general since its purpose was to guide and focus the early stages of knowledge elicitation without cultivating misinterpretations or premature conclusions. When the preliminary diagnostic model was shown to the experts, in diagrammatic form, they started adapting it to their specific problem domain with very little prompting. The new, more specific, model (figure 6) although more acceptable to the

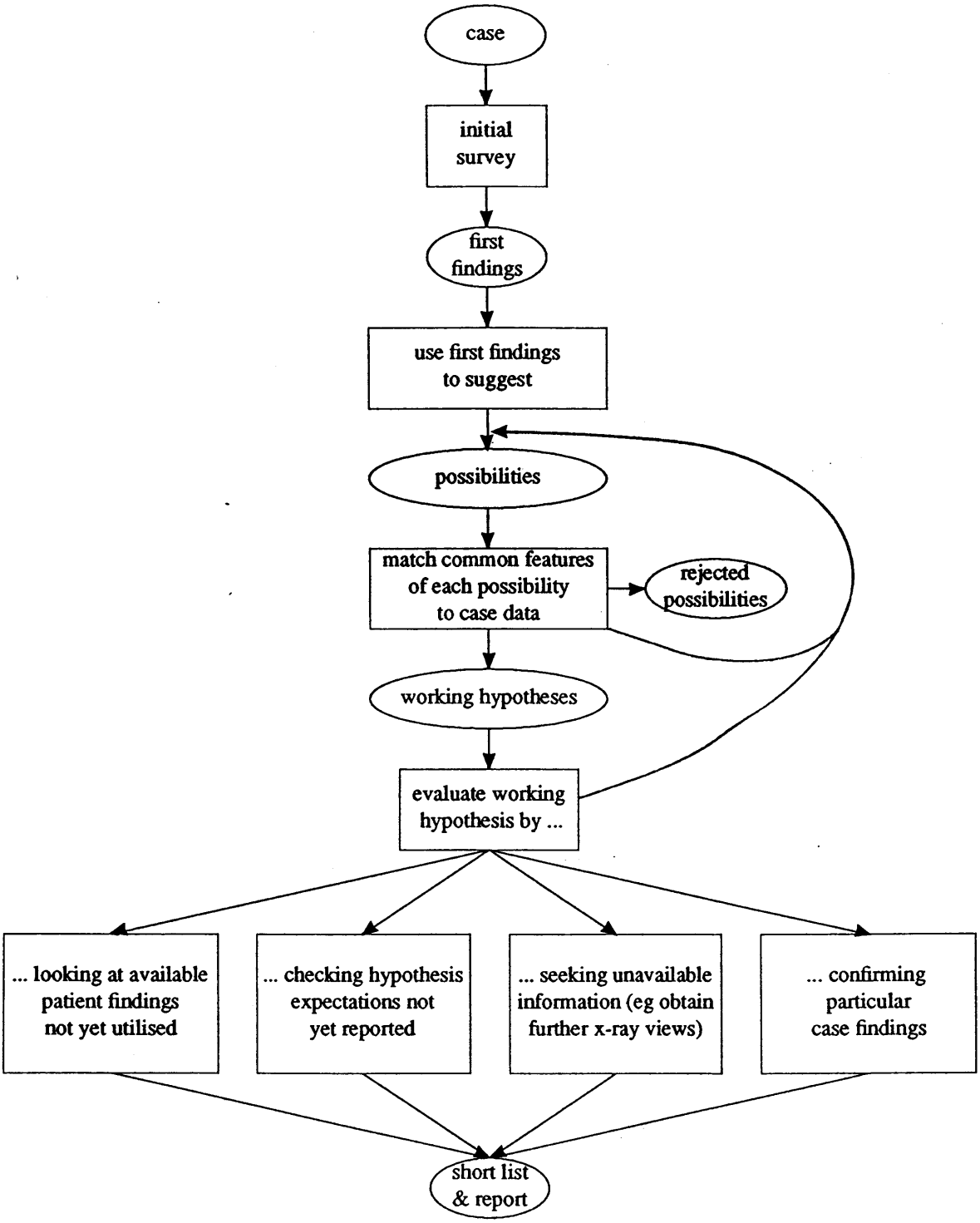


Figure 6 **Preliminary model of diagnostic reasoning.**

experts still needs to be further validated leading to more refinements. Much of the discussion was concerned with establishing a common terminology by attaching as a precise meaning as possible to the terms used. Once this was settled it became evident that some of the objections about the preliminary dysplasia model were due to misunderstandings about the meaning of terms.

The case findings consist of clinical data and X-ray images. The radiologist reads a skeletal survey and identifies chief radiological features, eg that there is a serious abnormality with the spine. Such striking features together with the clinical data generate certain possibilities about the case (abductive reasoning). Referring to the dysplasia model, primary triggers are instantiated suggesting possibilities. Often there are a few possibilities (around 5) and the next stage is to reject some of these. Each possibility is explored by checking whether the common features of the dysplasia fit the case findings; the radiologist may have to refer back to the X-ray images to check for more salient features. Matches on sufficient feature sets turn a possibility into a working hypothesis. Alternatively if the fit is not good enough the possibility is rejected. This reasoning stage is deductive in nature. The radiologist focuses on the common expectations of the triggered possibilities with a view to eliminating some of these possibilities and is not directly concerned with determining which case abnormalities are not accounted for.

Working hypotheses are seriously considered as final diagnoses. Normally the radiologist will be left with 2-3 working hypotheses which will be evaluated very closely. This is done by looking at available findings not yet utilised, checking hypothesis' expectations not yet observed (this may mean looking closer at available X-ray images), seeking currently unavailable information (eg obtaining further X-ray views), or confirming the accuracy of particular findings. This reasoning stage is largely inductive in nature; the radiologist wants to decide which working hypothesis provides the best explanation of the entire body of case findings.

When a possibility is being explored or a working hypothesis is being evaluated, new observations are expected to refer to features associated with the particular dysplasia. If "unexpected" findings (not supporting the pursued possibility or hypothesis) are obtained these may suggest another possible dysplasia (instantiating secondary triggers). Dysplasias associated through secondary triggers tend to share much of their features, which makes it likely for the presence of one to be confused with the presence of the other.

The current diagnostic model will provide the focus for the following stage of knowledge acquisition, which will lead to further refinements for the model. The model is complete in the sense that it currently includes a placeholder for every aspect of the diagnostic process; it is a skeleton for holding together contigencies about the diagnostic reasoning. One area that needs investigation is the refinement of hypotheses; for this dysplasia groups need to be analysed further.

The analysis of the diagnostic reasoning so far shows that the experts reason qualitatively. This is reflected in the dysplasia and feature models. Quantities were scarcely used, only for giving percentage estimations of the frequency of some

dysplasia manifestations. Hypotheses were evaluated qualitatively. Hypothesis evaluation is another area that needs further investigation. However, nothing in the current analysis points to the use of some numeric function for computing the overall "belief" in a hypothesis.

Hypothesis Status Transitions

Diagnosis involves the generation and evaluation of hypotheses, and this process can be modelled in terms of a graph called a Hypothesis Status-Transition Model (HSTM, figure 7). In the HSTM nodes represent possible status in the lifetime of a hypothesis and arcs represent transitions between status. A transition takes place when a condition is satisfied. The transition from *possibilities* to *working-hypotheses*, labelled "fit good enough", needs to be analysed further, as does the transition between working-hypotheses and the (complex) status *assessed.* The *suspended* status is a special status; conceptually it is seen as a pool for hypotheses which cannot be further progressed due to incomplete information. When this information becomes available, the suspended hypothesis reverts to its prior status (possibility or working hypothesis). Even if this information does not become available a suspended hypothesis can become assessed when other hypotheses related to it become assessed. In a diagnostic system which models hypothesis status and transitions, a rich justification for a hypothesis can be obtained by tracing the transitions in its status and their causes.

Focusing Through Triggers

Triggers provide an important focussing mechanism by generating possibilities. Primary triggers tend to be associated with groups of dysplasias, and this constrains the number of possibilities generated. Possibilities are also generated by secondary triggers during the evalution of another possibility or working hypothesis (as well as in the process of refining a dysplasia group).

As for hypotheses, the reasoning with primary triggers can be abstracted in terms of a status transition diagram (figure 8). A trigger is a set of one or more findings. When a trigger matches partially with case findings it becomes *potential.* If it is subsequently completely matched it becomes *valid* whilst if shown to be in conflict with case findings it becomes *invalid.* Valid triggers suggest possibilities and become *considered.* If a trigger cannot be completely matched due to incomplete information it becomes *suspended* until the information becomes available (if at all).

Hypothesis Evaluation

Currently a hypothesis that a dysplasia is present is generated on the basis of at least one match between sufficient feature sets for the dysplasia and case findings. The subsequent evaluation of hypotheses is very critical for the overall diagnostic performance. The analysis so far of this aspect of diagnostic reasoning for the problem domain does not suggest that the contending hypotheses are ranked on the basis of a single (numeric) estimation of their likelihoods. The reasoning is multi-dimensional in that hypotheses are ranked from different qualitative perspectives and there is no indication that these rankings are combined into an overall ranking. The following

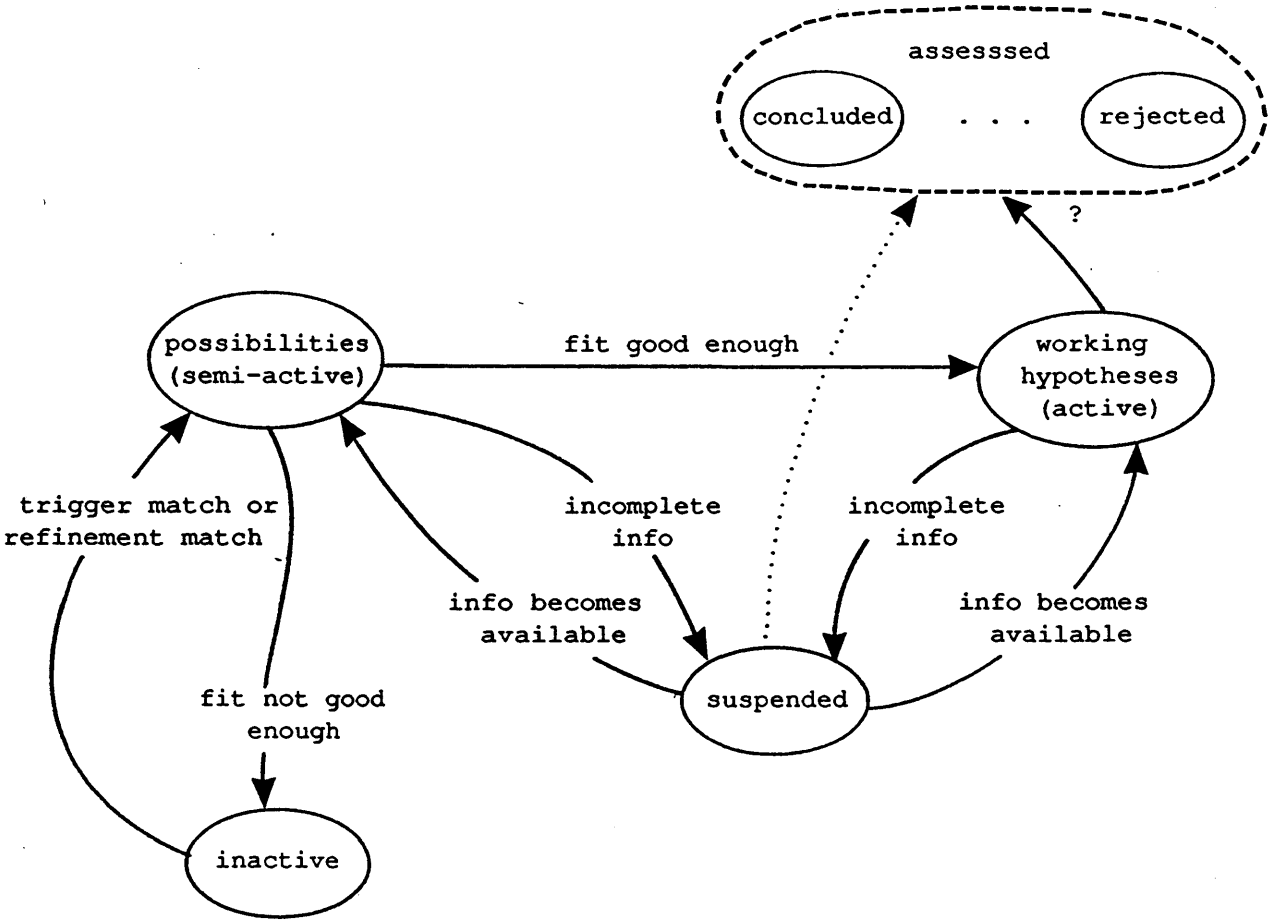


Figure 7 Hypothesis Status-Transition Model

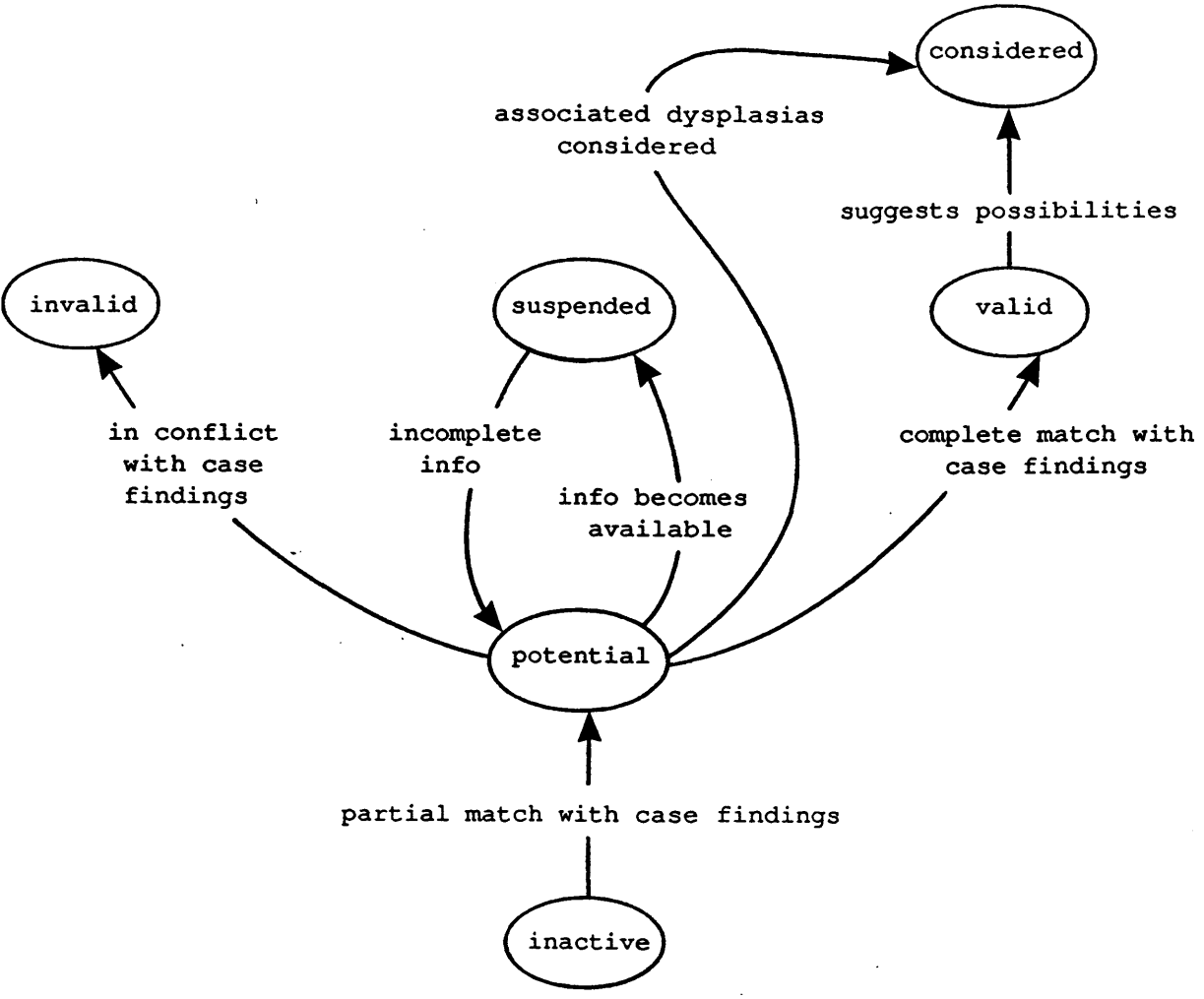


Figure 8 Primary Triggers Status-Transitions

perspectives have been identified:

- The proportion of the case findings matching common, typical or other features for the dysplasia.
- The proportion of case findings in conflict with common features of the dysplasia.
- The proportion of case findings which are irrelevant to the dysplasia.
- The proportion of the common features of the dysplasia in conflict with case findings.
- The proportion of the common features of the dysplasia in agreement with case findings.

The above indicate *how well* a hypothesis accounts for the case findings, and *how well* the hypothesis' expectations are met by the case findings. Most medical diagnostic systems, notably Internist-I (Miller et al, 1982), compute the match and mismatch between case findings and hypotheses' expectations; however they tend to merge this information into a single numeric value through a so-called *scoring function.*

5. SYSTEM ARCHITECTURE

The architecture of SDD (see figure 9) is based on the advanced architecture outlined in section 1. In this section the architecture is overviewed and a conceptual argument is provided for it. The architecture can also be justified on software engineering grounds but this is outside the scope of the paper.

Human experts not only have specific knowledge and expertise but also draw from a larger body of background knowledge. The latter forms foundational knowledge and aspects of it could be of a common-sense nature. Such knowledge may be called upon when the expert explains his decisions. A radiologist who is an expert on skeletal dysplasias will also be knowledgeable about bones in general (structural characteristics, concepts like ossification etc.) and familiar with clinical and other medical concepts outside radiology, although occasionally needing to consult specialists in these areas.

Background knowledge is not used by the diagnostic process directly. However the contribution of this knowledge to the overall problem solving activity is essential for competent behaviour (Keravnou and Johnson, 1987), its most important use being to "make sense" of the case findings so that the questions asked are intelligent. First generation expert systems lack background knowledge. The architecture of SDD alleviates such problems through the findings reasoner. The background knowledge is held in the finding subject frames.

Referring to figure 9 the diagnostic reasoning is distributed between a *hypotheses reasoner* (HR) and a *findings reasoner* (FR), each having its own knowledge base. The HR is the master requesting the services of the FR. This organisation is motivated by the MDX and PATREC systems (Mittal, 1980; Chandrasekaran and Mittal, 1983). The knowledge base for the HR holds the dysplasia frames. The knowledge base for the

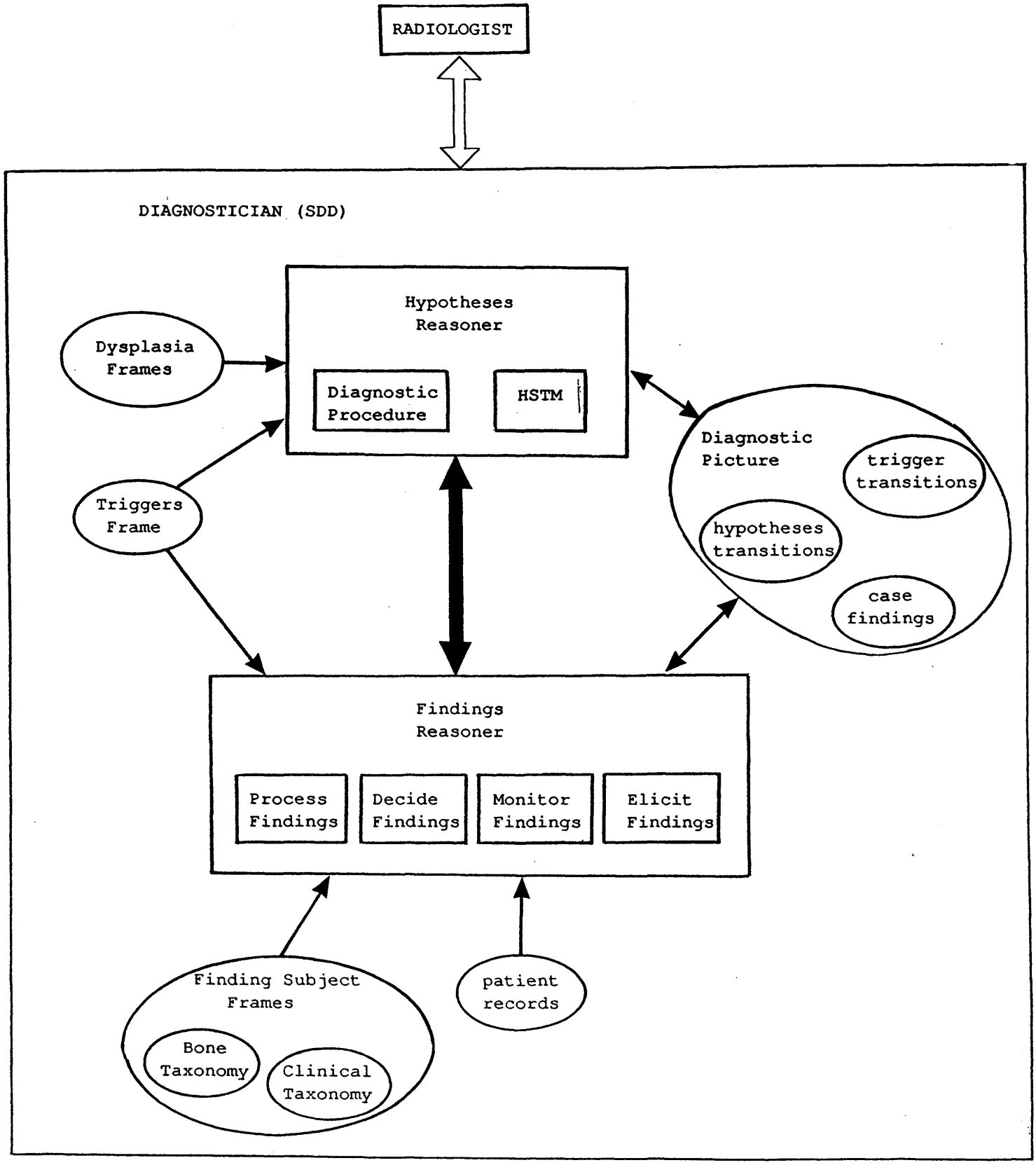


Figure 9 Overview of SDD Architecture

FR holds the finding subject frames. The FR also has access to patient records.

The case findings specific to the current consultation are kept in a global data structure, the *diagnostic picture*, which also keeps the hypotheses and their status transitions, and the triggers and their status transitions. The diagnostic picture is the placeholder for the operations of the HR and the FR. (A structure for a generic diagnostic picture is given in (Keravnou and Johnson, 1988)). New case findings are processed by the FR which checks that the findings are consistent (eg normal stature and short limbs are contradictory) and identifies potential primary triggers. The HR generates and evaluates hypotheses and decides which additional information to seek (eg for matching a trigger). The acquisition of new findings is guided by the FR. More specifically the functions of the FR are:

- To process new (user volunteered) case findings for consistency and to identify new potential primary triggers.
- To answer requests for information from the HR by deduction from the known case findings. (This is implemented in procedure Decide-Status (Keravnou and Johnson, 1987)).
- To guide the acquisition of additional findings required by the HR.
- To monitor the entry of new case findings (specified by the HR) and to inform the HR if such findings become true (this is used for instantiating secondary triggers and refinement suggestions).

The FR consists of a set of specific reasoners, eg a bone reasoner, a clinical reasoner, etc (figure 10). Its function is to deal with requests from the HR, which it does by delegating tasks to individual reasoners and collating their results. Each individual reasoner can perform the functions mentioned above but its access to the finding subjects knowledge base is restricted to those subjects relevant to it. The delegator decides which reasoner to invoke initially and deals with the result, which may be to invoke another reasoner. Having a reasoner invoke another reasoner indirectly through the delegator provides for a more rigid control structure. Consider the following example: Suppose the HR asks whether the limbs or spine are short. The delegator will invoke the bone reasoner. The bone reasoner sees that there is no explicit finding on limbs in the case findings, but knows that evidence of "short stature without dislocated femur" is sufficient to conclude short limbs or short spine. The case findings include Location of femur normal which enables the bone reasoner to establish that the femur are not dislocated. Limbs, femur and spine are in the domain of the bone reasoner. However stature is a clinical subject, outside of its domain. The bone reasoner passes the result of its operation to the delegator which then asks the clinical reasoner whether the stature is small. If the response from the clinical reasoner is positive the delegator will combine the two subresults and answer yes to the HR.

Figure 10 implies that there is a single level of reasoners. This can be extended to a multi-level situation where the intermediate reasoners are essentially delegators, as in the MDX system (Chandrasekaran et al, 1979).