Evidence-based Practice Center Systematic Review Protocol
Comparative Effectiveness of Treatments for Cryptorchidism

I. Background and Objectives for the Systematic Review

Cryptorchidism affects an estimated 3 percent of full-term male neonates and up to 30 percent of premature infants, making it the most common male genital anomaly identified at birth.\(^1\)\(^-\)\(^2\) It is a condition that is often apparent to parents and is easy to detect on routine physical examination. Therefore, boys with suspected cryptorchidism often present very early in life, often within the first year. Clinical decisionmaking about treatment is influenced by many factors, including whether or not the testicle is palpable, whether or not the condition is present unilaterally or bilaterally, the age at presentation, and coexisting medical conditions. In boys under 1 year of age whose testicle is palpable and is close to, but not quite inside, the scrotum, it may be difficult to distinguish between “true” cryptorchidism and a retractile testis. In this case, health care providers often elect to observe the patient’s condition until he is 1 year old.

Although about 70 percent of cryptorchid testes spontaneously descend within the first year of life, the number of boys whose condition persists remains constant at approximately 1 percent.\(^3\)\(^-\)\(^4\) Between 1992 and 2000, there were more than 600,000 (96 per 100,000) physician office visits for which cryptorchidism was the primary diagnosis.\(^1\) The annualized rate of orchiopexy was constant at 18 per 100,000 from 1994 to 1996.\(^5\) Given that the average submitted charges in 2008 for a level 3 in-office examination was $146 for new patients and $87 for established patients (according to physician payment information for value-driven health care data compiled by the Centers for Medicare & Medicaid Services) and that the costs of infant and postpubertal orchiopexy have been estimated to be $7,500 and $10,928, respectively, cryptorchidism incurs direct costs of millions of dollars annually, even by the most modest estimates.\(^6\) Cryptorchidism is, therefore, both a significant and costly health problem in the United States.

Treatment decisions may be guided by results of hormonal stimulation testing and/or imaging. The purpose of hormonal stimulation testing is to determine if viable testicular tissue is present. Specifically, if a boy has a nonpalpable testis, hormones such as human chorionic gonadotropins are administered to stimulate the testis. If increased levels of testosterone are noted after administration of human chorionic gonadotropins, it is assumed that there are viable testes somewhere in the body. If there is no response, the boy is usually presumed to be anorchid. Imaging is used to determine whether there is in fact a testis and to locate it. Imaging approaches include magnetic resonance angiography, which requires sedation or anesthesia and is thus not without immediate risks. In theory, absence of a testosterone increase in response to hormonal stimulation testing or inability to locate a testis with imaging should preclude the need for surgery as it indicates a lack of a potentially functional testis. However, the value and predictive power of these approaches for identifying the presence and location of a testis is currently not well understood, and their ability to prevent unnecessary surgery is an area of clinical uncertainty that is appropriate for systematic review.

Source: www.effectivehealthcare.ahrq.gov
Published Online: September 2, 2011
Associated conditions and consequences of cryptorchidism include hypospadias, hernia, and testicular torsion. Bilateral nonpalpable testes associated with hypospadias or ambiguous genitalia may represent severer developmental abnormalities (including intersexuality) that can be life threatening, warranting specific testing and treatment. Ascertaining the correct diagnostic workup for cryptorchidism is therefore important and includes identifying associated conditions.

Longer term consequences of cryptorchidism include testicular malignancy and infertility/subfertility, with stronger evidence for the etiologic role of cryptorchidism in malignancy than in disordered fertility. With regard to testicular cancer, it has been clearly established that there is a strong positive correlation between cryptorchidism and testicular cancer. An estimated 10 percent of all testicular tumors develop from an undescended testis, and the relative risk of incidence of a testicular tumor is about 40 times greater in men with cryptorchidism when compared to the general population.

For all outcomes, there remains clinical uncertainty in terms of selecting the optimal approach to workup (imaging vs. no imaging) and intervention (surgical vs. hormonal, orchiopexy, different surgical techniques). The ultimate goal of most interventions for cryptorchidism is to reposition the undescended gonad in a “normal” position in the scrotum. While there is some preliminary evidence that medical treatment with hormones, such as HCG, may result in descent of the cryptorchid testicle into the scrotum, the majority of interventions are surgical. The standard urology textbook, *Campbell-Walsh Urology*, considers that “early surgical repositioning of the testis into the scrotum before the onset of histopathological changes can reduce the risk of subfertility,” but this statement has not been systematically considered.

Although many clinicians advocate early orchiopexy to reduce the risk of testicular cancer associated with cryptorchidism, the published literature offers conflicting results.

Optimal timing of treatment for cryptorchidism is also uncertain, although there is some suggestion that performing surgical intervention early in life could reduce or eliminate the increased risk of testicular cancer.

Despite this uncertainty, there is growing clinical consensus that orchiopexy should be performed between 6 and 24 months of age.

II. The Key Questions

Introduction

We conferred with Key Informants familiar with the current state of the literature, clinical applications, and status of interventions for cryptorchidism in the United States in developing the Key Questions (KQs) and analytic framework. We distributed working drafts of the analytic framework and KQs prior to each discussion and solicited feedback on their utility, clarity, and relevance. Following these discussions, the KQs and framework were posted to the Agency for Healthcare Research and Quality’s Effective Health Care Program Web site for public comment for approximately 4 weeks. We received no comments during the public posting phase on the KQs or analytic framework.
Key Question 1a

For determining a course of treatment, is imaging equivalent to laparoscopy in determining the presence and location of a nonpalpable testicle?

Key Question 1b

In male children with bilateral, nonpalpable testes, does the use of hormonal stimulation testing reduce the need for surgery as part of a treatment plan?

PICOTS for Key Questions 1a and 1b

• Population(s):
  Prepubescent males presenting with cryptorchidism or suspected cryptorchidism

• Interventions:
  Workup evaluation for treatment planning, including imaging (magnetic resonance imaging, computerized tomography, and ultrasound), laparoscopy, and hormonal stimulation therapy; and hormones, including human chorionic gonadotropin or gonadotropin-releasing hormone

• Comparators:
  Other workup evaluation approaches for treatment planning (imaging, laparoscopy, hormonal stimulation therapy), nontreatment, later treatment, hormones

• Outcomes:
  1. Ability to correctly identify the presence and/or location of the testis
  2. Need for further surgical intervention
  3. Adverse effects, including but not limited to effects of sedation or anesthesia

• Timing:
  Timeframe for reporting of outcomes will not be restricted.

• Settings:
  All settings will be considered, including hospitals and university or academic medical centers. The effect of setting on diagnosis and therapy will be considered.
Key Question 2

What is the effectiveness of initial hormonal therapy (human chorionic gonadotropin or gonadotropin-releasing hormone) for the treatment of cryptorchidism for outcomes, including but not limited to:

- Further surgical intervention
- The effect on infertility/subfertility
- The development of testicular malignancy
- The size, location, and function of the testes

Key Question 3

What is the comparative effectiveness of surgical therapies (one-stage vs. two-stage, laparoscopic vs. open approach, orchiectomy vs. orchiopexy) for the treatment of cryptorchidism for outcomes including but not limited to:

- Further surgical intervention
- The effect on infertility/subfertility
- The development of testicular malignancy
- The size, location and function of the testes

Key Question 4

How does the age at presentation, physical presentation of cryptorchidism (unilateral vs. bilateral, palpable vs. nonpalpable, anatomic location) and occurrence of associated abnormalities (e.g., hernia) modify diagnosis, treatment, and outcomes?

Key Question 5

What is the nature and frequency of harms associated with workup or treatment for cryptorchidism?

PICOTS for Key Questions 2–5

- **Population(s):**
  
  Prepubescent males presenting with cryptorchidism or suspected cryptorchidism

- **Interventions:**

  Hormones including human chorionic gonadotropin or gonadotropin-releasing hormone; surgical therapy and specific surgical techniques (i.e., one-stage vs. two-stage orchiopexy, laparoscopic vs. open approach, orchiectomy vs. orchiopexy)
• **Comparators:**

Nontreatment, later treatment, hormones, and different surgical techniques

• **Outcomes:**

Immediate (within 6 weeks of therapy) and short-term (6 weeks to 2 years of therapy) outcomes:

1. Testicular size and appearance
2. Testicular position
3. Pain
4. Parent/patient satisfaction
5. Need for further surgical intervention
6. Emotional/psychosocial response
7. Adverse effects, including but not limited to pain, infection, hematoma, and edema

Long-term (more than 2 years after therapy) outcomes:

1. Testicular size and appearance
2. Testicular position
3. Endocrine function
4. Body image
5. Parent/patient satisfaction
6. Infertility/subfertility
7. Torsion
8. Testicular malignancy and cancer
9. Hernia
10. Emotional/psychosocial response

• **Timing:**

Timeframe for reporting of outcomes will not be restricted.

• **Settings:**

All settings will be considered, including hospitals and university or academic medical centers. The effect of setting on diagnosis and therapy will be considered.
III. Analytic Framework

Input from the Key Informants was crucial in shaping the analytic framework. No comments related to the framework were received during the public posting phase.

Figure 1. Analytic framework for treatments of cryptorchidism

*Initial hormone therapy versus immediate surgical therapy or surgical therapy versus surgical therapy.

UDT = undescended testicle

IV. Methods

A. Criteria for Inclusion/Exclusion of Studies in the Review –

Table 1 lists the inclusion and exclusion criteria developed during the topic refinement phase. The criteria are based on our understanding of the current literature, input from content experts and Key Informants, and established principles of methodological quality.

We set a cut-off year of 1980 for all included publications. This date was decided on during discussions with our Key Informants and content experts and was incorporated as a limit into the initial literature search. Currently, computerized tomographic imaging, ultrasound, and laparoscopy are commonly employed in the workup and treatment of cryptorchidism. These technologies diffused into widespread use in the 1980s. Reports prior to this time period are not clinically relevant to today's patients. To this end, our literature search will be limited to
publications dated 1980 or later. Our Key Informants and content experts believe that this cut-off will not affect the quality or applicability of any findings.

We will focus the review on studies published in English; included studies may include non-U.S. populations but must be published in English. Our Key Informants and content experts believe that exclusion of non–English-language studies will not affect the quality or applicability of any findings, as the vast majority of studies will be published in English. This language limit was also incorporated into the initial literature search. Further inclusion and exclusion criteria are described below.

Table 1. Inclusion and exclusion criteria

<table>
<thead>
<tr>
<th>Category</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study population</td>
<td>Prepubescent males presenting with cryptorchidism or suspected cryptorchidism</td>
</tr>
<tr>
<td>Time period</td>
<td>Studies published after 1980</td>
</tr>
<tr>
<td>Publication languages</td>
<td>English only</td>
</tr>
<tr>
<td>Admissible evidence (study design and other criteria)</td>
<td>Admissible designs</td>
</tr>
<tr>
<td></td>
<td>• All study designs will be considered, except case reports</td>
</tr>
<tr>
<td></td>
<td>• For KQ1 and KQ2, single-arm studies may be included</td>
</tr>
<tr>
<td></td>
<td>• For KQ3 and KQ4, studies must use a relevant comparison group (i.e., comparison of different treatments, hormonal vs. surgical therapy, treatment vs. no treatment)</td>
</tr>
<tr>
<td>Other criteria</td>
<td>• Original research studies that provide sufficient detail regarding methods and results to enable use and adjustment of the data and results</td>
</tr>
<tr>
<td></td>
<td>• Studies detailing only the prevalence or etiology of cryptorchidism will not be included (i.e., an intervention for cryptorchidism must be included)</td>
</tr>
<tr>
<td></td>
<td>• Studies of men presenting with other conditions, such as infertility or testicular cancer, that work backwards to determine whether past treatment for cryptorchidism may be a cause will not be included due to the inherent difficulty in accurately capturing information on past procedures</td>
</tr>
<tr>
<td></td>
<td>• Studies that include intersex individuals or individuals with ambiguous genitalia will not be included, even if participants undergo treatment for cryptorchidism, as this treatment is likely to be in concert with treatment for associated conditions that may compound the difficulty in drawing meaningful conclusions</td>
</tr>
<tr>
<td></td>
<td>• Studies must include at least one outcome measure of an outcome listed in the PICOTS</td>
</tr>
<tr>
<td></td>
<td>• Relevant outcomes must be extractable from data presented in the papers</td>
</tr>
</tbody>
</table>
B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies To Answer the Key Questions

Search the Literature

To ensure comprehensive retrieval of relevant research into the treatments for cryptorchidism, our approach to the literature will include three key databases: the PubMed®, medical literature database, the Cumulative Index of Nursing and Allied Health Literature (CINAHL®), and the EMBASE® Drugs & Pharmacology database. Search strategies in each of these databases will focus specifically on terms related to treatment for cryptorchidism, including keywords and subject terms representing cryptorchidism interventions and excluding non–English-language materials, literature related to nonhuman subjects, and publications not resulting from some form of clinical trial (e.g., reviews, letters, commentaries, and others).

We will update the search quarterly during the abstract and full-text review stages, adding relevant references to the pool of articles under consideration as needed. We will also update the search after submission of the draft report and add relevant references as needed while the draft report is undergoing review. We will also incorporate references that meet our inclusion criteria or are of particular relevance for background sections that may be brought forward by public/peer reviewers.

We will employ additional searches of the reference lists of recent existing systematic reviews or meta-analyses of treatments for cryptorchidism; the investigative team will also scan the reference lists of articles undergoing full-text review for citations potentially meeting our inclusion criteria.

Search for Grey Literature

We will conduct a broad search for grey literature relevant to the topic of our review, including meeting abstracts, conference proceedings, and reports. We will also seek suggestions from the Technical Expert Panel (TEP) with regard to additional potential sources of grey literature. We will incorporate relevant information from grey literature searches into the review as appropriate (i.e., for assessing publication bias or selective outcomes reporting).
Develop Abstract and Full-text Review Forms

We will develop separate forms for the abstract and full-text reviews. The abstract review forms will contain questions about our primary exclusion/inclusion criteria. The full-text review forms will be more detailed and are intended to assist in a) identifying studies that meet the inclusion criteria and b) conducting an initial sort of studies into the appropriate KQs. Finally, data-extraction forms will be used to collect those data necessary for evidence tables and synthesis.

Initial Review of Abstracts

We will review all titles and abstracts identified through searches against our inclusion/exclusion criteria. Each abstract will be reviewed by at least two members of the investigative team. When differences between the reviewers arise, we will err on the side of inclusion. For studies without adequate information to make the determination, we will retrieve the full articles and review them against the inclusion/exclusion criteria.

Retrieve and Review Articles

We will retrieve and review all articles meeting our predetermined inclusion/exclusion criteria or for which we have insufficient information at the abstract phase to make a determination about eligibility. Each article will be reviewed by two members of the investigative team. When differences between the reviewers arise, a third reviewer will adjudicate the results.

C. Data Extraction and Data Management

Determine Outcomes To Extract

Specific outcomes to extract have been identified a priori during conversations with our TEP and are presented in the PICOTS section of this document. Primary outcomes include testicular size, position, and appearance; parent/patient satisfaction; psychosocial and emotional response; and the need for further intervention. Outcomes may occur in both the short and long term. Additional outcomes may include, but are not limited to, pain, testicular function, body image, fertility, torsion, cancer, and hernia. Adverse outcomes (harms) will also be captured and discussed.

The feasibility of extracting outcomes is dependent on the quality of available literature. The proposed outcomes to extract have been determined by our internal team and TEP. Outcomes to extract may change based on the review of full-text articles meeting the inclusion criteria, at which point a protocol amendment will be completed if necessary.

For the studies meeting the conditions of the second-round assessment, the abstractors will extract key data and study quality elements from the article(s) and enter them into evidence tables. The Methods and Content Leads and content experts will review the extraction forms.
against the original articles for quality control. Differences between the abstractor and the reviewer will be resolved by consensus.

We will develop a simple categorization scheme for coding the reasons that articles, at the stage of full review, are not finally included in the report. The abstractor will note the reason for exclusion on the article cover page. We will then record that code in an EndNote® database, our bibliographic software, so that we can later compile a listing of excluded articles and the reasons for such exclusions.

D. Assessment of Methodological Quality of Individual Studies

Assess Study Quality

The quality of individual studies will be assessed by using specific, established assessment tools for each type of study. For randomized controlled trials, the Cochrane Collaboration’s tool for assessing risk of bias will be employed.\textsuperscript{14} Fundamental domains will include: adequate sequence generation, allocation concealment, blinding, incomplete outcome data addressed, and free of selective reporting bias.

For nonrandomized and observational studies, the Newcastle-Ottawa scale will be used.\textsuperscript{15} The scale assesses three broad perspectives: 1) the selection of the study groups; 2) the comparability of the groups; and 3) the ascertainment of either the exposure or outcome of interest for case-control or cohort studies, respectively. For example, for a cohort study, the fundamental criteria will include: representativeness of cohort, selection of nonexposed cohort, ascertainment of exposure, outcome of interest, comparability of cohorts, assessment of outcome, adequate duration of followup, and adequate followup of cohort. Other sources of bias can include baseline imbalances, source of funding, early stopping for benefit, and appropriateness of crossover design.

For diagnostic studies, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool will be used.\textsuperscript{16} Items include representativeness of patient spectrum, selection criteria, reference standard, disease progression bias, verification bias, review bias, test execution, study withdrawals, and uninterpretable or indeterminate results.

We will use a components approach to assess and report quality. Two senior staff will independently perform quality assessment of the included studies with disagreements resolved through discussion or third-party adjudication as needed. We will record quality assessments in tables, summarizing for each study.

E. Data Synthesis

Prepare Evidence Tables

We will enter data into evidence tables, using predetermined abbreviations and acronyms and otherwise attending to consistency across entries from the outset. The dimensions (i.e., areas of special focus, or the columns) of each evidence table may vary by KQ as appropriate, but the tables will contain some common elements, such as author, year of publication, study location (e.g., country, city, state) and time period, population description, sample size, and study type (e.g., randomized controlled trial, prospective observational study, etc.).

\textbf{Source:} \texttt{www.effectivehealthcare.ahrq.gov}

\textbf{Published Online:} September 2, 2011
F. Grading the Evidence for Each Key Question

Assess the Strength of Evidence

We will also use explicit criteria for rating the overall strength of the collective evidence on each KQ into qualitative categories (e.g., low, moderate, high, insufficient). We will use established concepts of the quantity of evidence (e.g., numbers of studies, aggregate ending-sample sizes), the quality of evidence (from the quality ratings on individual articles), and the coherence or consistency of findings across similar and dissimilar studies and in comparison to known or theoretically sound ideas of clinical or behavioral knowledge. We will make these judgments for each of the main KQs and any subquestions related to specific outcomes as appropriate.

The strength-of-evidence evaluation will be that stipulated in the Agency for Healthcare Research and Quality’s Methods Guide for Effectiveness and Comparative Effectiveness Reviews, which emphasizes the following four major domains: risk of bias (low, medium, high), consistency (inconsistency not present, inconsistency present, unknown or not applicable), directness (direct, indirect), and precision (precise, imprecise). Risk of bias is derived from the quality assessment of the individual studies which addressed that KQ and specific outcome under consideration. Each key outcome on each comparison of interest will be given an overall evidence grade based on the ratings for the individual domains.

The overall strength of evidence will be graded as “high” (indicating high confidence that the evidence reflects the true effect and further research is very unlikely to change our confidence in the estimate of effect); “moderate” (indicating moderate confidence that the evidence reflects the true effect and further research may change our confidence in the estimate of effect and may change the estimate); “low” (indicating low confidence that the evidence reflects the true effect and further research is likely to change our confidence in the estimate of effect and is likely to change the estimate); or “insufficient” (indicating that evidence is either unavailable or does not permit estimation of an effect). When no studies are available for an outcome or comparison of interest, the evidence will be graded as insufficient.

Two senior staff will independently grade the body of evidence, with disagreements resolved through discussion or third-party adjudication as needed. We will record strength of evidence assessments in tables, summarizing for each outcome.
G. Assessing Applicability

Our team will assess the applicability of findings reported in the included literature to the general population of adolescents with cryptorchidism by determining the population, interventions, comparators, timing, and setting in each study and developing an overview of these elements. This assessment will be done to account for any factors limiting the ability to apply the intervention to other populations or other settings, such as inadequate description of the intervention or failure to report critical data. We will also review potential modifiers of effect of treatment, including age at presentation, medical history, overall health and other medical conditions, and sidedness (unilateral vs. bilateral).

V. References

VI. Definition of Terms

The following definitions are taken from *Campbell-Walsh Urology*:

*Laparoscopy*: minimally invasive diagnostic surgery that can be used to identify the presence or absence, the location, and the anatomy of the nonpalpable testis

*Orchiopexy*: surgical placement of an undescended testis in the scrotum

*Orchietomy*: surgical removal of the testis

VII. Summary of Protocol Amendments

In the event of protocol amendments, the date of each amendment will be accompanied by a description of the change and the rationale.

VIII. Review of Key Questions

For all EPC reviews, key questions were reviewed and refined as needed by the EPC with input from Key Informants and the TEP to assure that the questions are specific and explicit about what information is being reviewed. In addition, for Comparative Effectiveness reviews, the key questions were posted for public comment and finalized by the EPC after review of the comments.

IX. Key Informants
Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts comprise a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for CERs and Technical briefs, be published three months after the publication of the Evidence report.
Potential Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.