

Commentary on the articles by Collet, You et al., and Kim et al.

Das ist die Tragödie der Wissenschaft: Hässliche Fakten zerstören schöne Hypothesen

Als Viggo Andresen und Karl Häupl Mitte der dreißiger Jahre ihr „norwegisches System der Funktionskieferorthopädie“ in das junge Fach der Kieferorthopädie einführten, lösten sie damit in Deutschland eine veritable fachliche Revolution aus. Endlich, so glaubten sie, stünde ein über das mechanische Verschieben von Zähnen weit hinausgehendes, fundamental neues Therapieprinzip zur Verfügung, mit dessen Hilfe das Wachstum des Unterkiefers oder seine Lage zum Schädel verändert werden könnte. Die Wirkung des Aktivators wurde als passiv und funktionell den festsitzenden, mechanisch wirkenden Geräten gegenübergestellt.

Zumindest in Deutschland und einigen Nachbarländern wurde die Funktionskieferorthopädie für viele Jahrzehnte zum vorherrschenden Therapieprinzip bei der Behandlung der Distal- und Mesialokklusion, während die festsitzenden Apparaturen bis in die siebziger Jahre hinein als zweitklassig, mechanistisch und obsolet galten.

Und das alles, diese großartige lange Tradition, zahllose wissenschaftliche Arbeiten, der Kernbestand fachlichen Wissens von Generationen deutscher Kieferorthopäden, soll jetzt falsch sein? Haben sich Andresen und Häupl, Petrik, Eschler, Korkhaus, Hausser und viele andere der Großen unseres Fachs einfach geirrt?

Sehen wir uns die Arbeiten aus dem deutschen Sprachraum einmal näher an, mit denen die skelettale Wirkung der

propulsiv wirkenden, sogenannten FKO-Geräte untermauert werden sollte. So fällt zunächst auf, dass schon Andresen und Häupl in ihren ersten Publikationen von verlängerten oder verlagerten Unterkiefern ausgingen, obwohl sie ihre Beobachtungen ausschließlich an Kiefermodellen anstellten. Das seitliche Fernröntgen war zwar bereits 1934 von Hofrath in Deutschland publiziert worden, gehörte aber bis in die fünfziger und sechziger Jahre hinein in der deutschen Kieferorthopädie nicht zum diagnostischen Standardrepertoire, so dass in den „Fortschritten der Kieferorthopädie“ auch um 1960 noch das Modellfoto dominierte. Allein mit Hilfe des FRS hätte jedoch gemessen werden können, ob Zahn- oder Kieferbewegungen zum Therapieresultat beigetragen haben. Es kommt hinzu, dass die meist vor 1950 publizierten FRS-Analysen nur bedingt geeignet waren, skelettale und dentale Veränderungen mit hinreichender Genauigkeit zu unterscheiden. Dieser Mangel führte erst in späteren Jahrzehnten zur Entwicklung speziell für diese Fragestellung zugeschnittener Analysen wie der strukturellen Überlagerung nach Björk, der Pancherz-Analyse und der Johnston-Pitchfork-Analyse. Das wichtigste Axiom der FKO, nämlich ihr vorgebliches Potenzial, dauerhaft das Schädel- und Gesichtswachstum verändern zu können, war aber bereits lange vorher formuliert worden.

Sowenig wie das röntgenologische Messbesteck genügt die Methodik nahezu aller älteren Arbeiten den Ansprüchen wissenschaftlicher Beweisführung. Vorherrschender Typus ist lange Jahre der Fallbericht, und erst nach und nach setzt sich die retrospektive Untersuchung größerer Gruppen mit statistischer Auswertung durch. Mit Recht weist Collett aber

Korrespondenzadresse

Dr. Henning Madsen · Ludwigstr. 36 · 67059 Ludwigshafen · T 0621-59 16 80 F 0621-59 16 820
e-mail: info@madsen.de · <http://www.madsen.de>

Bibliografie

Inf Orthod Kieferorthop 2002; 34: 265 – 268 © Georg Thieme Verlag KG · ISSN 0020-0336

darauf hin, dass auch die retrospektive Untersuchung eine Methodik mit grundsätzlichen Schwächen ist. Wie werden denn die untersuchten Gruppen zusammengestellt? In der Regel doch so, dass aus der Patientenkartei Patienten mit erfolgreich abgeschlossener Behandlung ausgesucht werden. Dabei fallen alle diejenigen weg, deren Behandlung nach längerer Erfolglosigkeit umgestellt wurde, die aus Unzufriedenheit oder anderen Gründen abgebrochen haben oder deren Behandlung mit unbefriedigendem Ergebnis abgeschlossen wurde.

Ausgerechnet Rolf Fränkel hat das Elend der retrospektiven Untersuchung in einer abschließenden Betrachtung zu seinen eigenen Untersuchungen in dankenswerter Offenheit auf den Punkt gebracht, als er ausführte „only patients were included who had an unstrained competent lip seal at the end of retention“ [3]. Korrekt müssten derartige Arbeiten etwa als „Nachuntersuchung der schönsten Fälle der xy-Therapie“ titulierte und es dürften keine allgemeingültigen Aussagen aus ihnen abgeleitet werden.

Die Aussagekraft zahlreicher retrospektiver Untersuchungen wird durch das Fehlen einer geeigneten Kontrollgruppe weiter verschlechtert, so dass in vielen Fällen offensichtlich nichts weiter als das natürliche Wachstum der Patienten als skelettaler Therapieeffekt gewertet wurde.

Der Einwand, es seien Tierversuche mit beeindruckender Beweiskraft für die skelettalen Effekte der propulsiven Geräte publiziert worden, ist nicht stichhaltig. Nahezu alle Tierversuche sind an Tieren mit einer kurzen Wachstumsperiode von wenigen Wochen bis Monaten durchgeführt worden, meist an Nagetieren, seltener an kleinen Affen. Bei diesen Tieren ist es leicht möglich, mit einer sehr kurzen Versuchszeit einen großen Teil der gesamten Wachstumsperiode abzudecken. Beim Menschen müssten Therapiezeiten von 10–15 Jahren veranschlagt werden, um die in den Tierversuchen erzielten, skelettalen Effekte zu reproduzieren, was aus vielen Gründen nicht möglich ist.

Nur am Rande sei bemerkt, dass das FKO-Prinzip bei den Versuchstieren dahingehend abgewandelt wurde, dass feststehende Vorschubgeräte zum Einsatz kamen, womit auch die „Passivität“ und das „Funktionelle“ dieser Apparate verloren ist. Diese simplen Fakten zeigen, dass die nahtlose Übertragung der Ergebnisse der bisherigen Tierversuche auf den Menschen prinzipiell nicht möglich ist.

Das Untersuchungsprotokoll mit der höchsten Aussagekraft ist die randomisierte, kontrollierte Studie (RCT), bei der vor der Untersuchung die Patienten nach Zufall auf die Therapie- und die Kontrollgruppe verteilt werden. Durch die Zufallsverteilung wird vermieden, dass beispielsweise Patienten mit bestimmten, günstigen Merkmalen bevorzugt der Therapiegruppe zugeordnet werden. Die gleichzeitig beobachtete, statistisch vergleichbare Kontrollgruppe erlaubt es, therapeutische Effekte von solchen Effekten zu unterscheiden, die durch natürliches Wachstum oder durch andere, bekannte Standardtherapien hervorgerufen werden.

Da die Ergebnisse von unveränderten, zufällig entstandenen Gruppen ausgewertet werden, ermöglicht der RCT am ehesten von allen Untersuchungsprotokollen Aussagen über zukünftige, ebenfalls zufällig zusammengesetzte Gruppen, wie z. B. die nächsten 100 Patienten in unserer Praxis. Leider ist es wesentlich aufwändiger, einen RCT durchzuführen, als eine retrospektive Untersuchung, was die Zahl der möglichen RCTs empfindlich limitiert. Nichtsdestotrotz gibt es inzwischen eine Reihe RCTs zur Funktionskieferorthopädie, die von Collett meines Wissens vollständig aufgelistet werden.

Diese RCTs zeigen fast durchweg das gleiche Bild: wenn überhaupt statistisch signifikante Unterschiede zwischen Patienten mit FKO-Geräten und solchen mit Headgear bzw. Multibracketapparatur festgestellt werden, so sind diese so klein, dass sie ohne klinische Relevanz sind. Bessere kurzfristige Resultate finden sich bei festsitzenden (Herbst) oder nahezu auf Vollzeitbasis getragenen herausnehmbaren Apparaturen (Twinblock), jedoch darf auch bei diesen nicht mit posttherapeutischer Stabilität der skelettalen Veränderungen gerechnet werden.

Werden bessere langfristige Resultate berichtet, so ist dies fast immer auf methodische Mängel zurückzuführen*. Und selbst wer die um 0,72 mm längeren Unterkiefer in der FKO-Gruppe, die Tulloch et al. [9] fanden, für klinisch bedeutsam hält, wird ernüchert feststellen müssen, dass auch diese kleine Veränderung nicht stabil bleibt.

Die Langzeituntersuchungen nach abgeschlossener FKO-Therapie weisen fast einstimmig darauf hin, dass die direkt nach Therapieabschluss gemessenen skelettalen Effekte in den Folgejahren bis zur statistischen Insignifikanz zu verschwinden pflegen. In den auf die aktive Behandlung folgenden Jahren weisen FKO-Patienten erheblich weniger Unterkieferwachstum auf als unbehandelte oder konventionell behandelte Vergleichsgruppen, so dass das ursprünglich vorhandene skelettale Muster sich weitgehend wiederherstellt.

Alle weiteren angeblichen Vorzüge der FKO, etwa die Profilverbesserung, die Vermeidung von Extraktionen und Dysgnathieoperationen sind damit hinfällig, abgesehen davon, dass die meisten dieser Detailfragen ebenfalls schon mit negativem Ausgang für die FKO untersucht wurden. Trotzdem ist es unstrittig, dass es in vielen Fällen möglich ist, eine Distalokklusion mit einem FKO-Gerät in eine Neutralokklusion zu überführen.

* Ein Beispiel für gravierende methodische Mängel bietet die Arbeit von Mills & McCulloch [7]. Die Autoren behaupten, die posttherapeutische Stabilität von Wachstumsveränderungen des Unterkiefers, die mit dem Twinblock erzielt wurden, über drei Jahre nach Therapieende beobachtet zu haben. Liest man die Arbeit jedoch genau, so findet sich, dass die Twinblocks posttherapeutisch durchschnittlich 18 Monate weiter nachts getragen wurden, tatsächlich also rund die Hälfte der Nachbeobachtungszeit in Wirklichkeit nur eine verlängerte Therapiezeit war. Dadurch erklärt sich auch die in anderen Langzeituntersuchungen nicht festgestellte Persistenz der Wachstumsinkremente.

Die Arbeiten von You und Kim geben wichtige Hinweise, warum das so ist – auch ohne Wachstumsstimulation. In der Kieferorthopädie wurde seit den Untersuchungen von Brodie vermutet, dass der bestehende Gesichtstyp auf das Ausmaß des zukünftigen Wachstums schließen lasse: Dieser Annahme nach war also bei der Klasse I mit viel, bei der Klasse II mit wenig Unterkieferwachstum zu rechnen. Dies wurde von US-Kieferorthopäden in der griffigen Formulierung „them that has, gets“ (wer hat, der bekommt) zusammengefasst.

Die Untersuchung von You zeigt dagegen, dass diese Auffassung falsch ist. Es ist das angeborene Recht der Klasse-II-Patienten, durchschnittlich 4,36 mm mehr Wachstum im Unterkiefer als im Oberkiefer zu erleben – ein Betrag, der rund zwei Drittel einer mit 7 mm veranschlagten vollen Klasse II ausmacht und keiner therapeutischen Stimulation bedarf. Diese Angabe ist eine bemerkenswert genaue Reproduktion des in einer früheren Studie (Karlsen & Krogstad, 1999) festgestellten Wertes von 4,6 mm, so dass hier von einem sehr hohen Evidenzgrad ausgegangen werden kann. Trotz dieses für die Therapie so günstigen Wachstums korrigiert sich eine vollständige Klasse-II-Verzahnung natürlich nicht von selbst, und in den Arbeiten von Kim und You wird sehr präzise gezeigt, warum das nicht geschieht: Im selben Umfang, den der Unterkiefer den Oberkiefer im Wachstum übertrifft, verlagert sich die obere Dentition nach vorn, die untere nach hinten, um die Verhöckerung der Zähne aufrecht zu erhalten. Beide Autoren kommen letztendlich zu dem Schluss, dass das Prinzip der Klasse-II-Therapie immer das gleiche ist: Durch die Entschlüsselung der Okklusion wird die dentoalveoläre Kompensation des überschießenden Unterkieferwachstums verhindert, so dass dieses therapeutisch ausgenutzt werden kann. Es spielt dabei keine Rolle, ob die Verschlüsselung durch Klasse-II-Gummizüge, Headgear, eine Aufbissplatte oder einen Bionator aufgehoben wird.

Und damit schließt sich der Kreis, denn so ist es auch zwanglos zu erklären, dass die meisten vergleichenden Untersuchungen keine relevanten Unterschiede zwischen FKO und Mechanotherapie finden konnten: Gleiche Therapieprinzipien erzielen einfach gleiche Ergebnisse.

Nun ist mir durchaus bewusst, dass diese Sätze manchem gestandenen Kollegen als weit hergeholt, gar provokativ erscheinen mögen. Einwände sollten jedoch in jedem Fall mit entsprechender Literatur untermauert werden. Die Übersichtsarbeit von Collett steht im Übrigen nicht allein; etliche weitere Reviews mit ähnlichem Resultat sind seit Jahren veröffentlicht, deren beste wahrscheinlich Aelbers & Dermant [1] ist, die systematischer als Colletts Untersuchung angelegt ist und auch noch eine quantitative Metaanalyse der veröffentlichten Arbeiten unternimmt. Sehr lesenswert ist auch Johnston [5]. Allen diesen Arbeiten ist gemeinsam, dass sie ohne Widerhall in der deutschen Kieferorthopädie geblieben sind, also kaum jemand sie wahrgenommen hat. Amerikanische Unterkiefer wachsen aber nicht anders als deutsche, die „scientific community“ ist – von einigen weißen Flecken abgesehen – weltumspannend, und daher sollte der ernüchternde Forschungsstand zur Klasse-II-Therapie auch in Deutschland zur Kenntnis genommen werden.

Die Quintessenz: von den propulsiv wirkenden kieferorthopädischen Geräte darf im Durchschnitt nur ein geringer skelettaler Effekt in der Größenordnung eines Millimeters erwartet werden. Selbst dieser klinisch kaum relevante Effekt ist über die Folgezeit nicht stabil, so dass er in der Regel einige Jahre nach Behandlungsende nicht mehr wahrzunehmen ist: Es scheint, als ob sich diese Geräte günstigstenfalls etwas Wachstum aus der Zukunft leihen würden, das in der Folgezeit auf Cent und Euro „zurückgezahlt“ werden muss.

Apparate wie das Herbstscharnier mitsamt seinen Klonen, wie auch auf Vollzeitbasis getragene herausnehmbare Geräte haben bessere initiale Effekte, jedoch scheint danach die posttherapeutische „stille Periode“ des Unterkieferwachstums besonders ausgeprägt zu sein. Gerade das effizienteste propulsive Gerät, das Herbstscharnier, basiert auf einer mechanischen Verblockung der Zahnreihen und nicht auf Muskelfunktion. Aus diesen Gründen kann bei den propulsiven Geräten nicht von einem von der Mechanotherapie unterschiedlichen Therapieprinzip gesprochen werden, das auf Muskelfunktion basiert, denn entscheidend ist bei den propulsiven Geräten nicht die Kraftquelle, sondern die therapeutische Vorverlagerung des Unterkiefers. Der Begriff der Funktionskieferorthopädie sollte daher durch propulsive Therapie ersetzt und die Ideen der „Passivität“ und des „Funktionellen“ dieser Apparaturen aufgegeben werden. Sie sind letzten Endes orthodontisch wirkende Apparate, nicht mehr und nicht weniger. Für ihren Einsatz sollten künftig Effizienzkriterien herangezogen werden. So zeitigt das in Deutschland bis heute übliche Konzept, zunächst mit einer FKO-Phase die Okklusion zu korrigieren, um anschließend in einer zweiten Phase mit MB-Apparatur weiterzuarbeiten, keineswegs bessere Ergebnisse, sondern nachweislich längere Behandlungszeiten und höhere Kosten [4, 6, 8, 9]. Der Großteil unserer Patienten kann aber in einer einzigen Phase mit einer MB-Apparatur in maximal 24 Monaten behandelt werden, was vermutlich eine Halbierung der hierzulande üblichen Behandlungszeiten bedeuten würde.

Wer trotzdem auf zwei Behandlungsphasen und propulsive Geräte nicht verzichten mag, sollte seinen Patienten zumindest die Möglichkeit einer wesentlich kürzeren Einphasenbehandlung anbieten und ihnen die Wahl überlassen. So wie sich die jahrzehntelang gepflegte Auffassung, Gesichts- und Schädelwachstum stünden unter der Herrschaft von „kleinen, eponymen Stücken Plastik“ (Johnston) als Überschätzung erweist, darf heute mit Fug und Recht auch die Umbenennung unseres Fachs vom ursprünglichen Begriff „Orthodontie“ in „Kieferorthopädie“ in Frage gestellt werden. Geradezu eine Ironie der Geschichte ist es vor diesem Hintergrund, dass ausgerechnet die amerikanische Fachgesellschaft AAO Mitte der achtziger Jahre der bei weitem wichtigsten Fachzeitschrift der Welt, dem „American Journal of Orthodontics“, den Zusatz „and Dentofacial Orthopedics“ verordnete. Dentofacial orthopedics? – Sorry, there is no such thing.

Die relevante Literatur ist zum großen Teil im Literaturverzeichnis von Collett aufgeführt. Weitere Literaturstellen finden sich in den Literaturverzeichnissen von Aelbers &

Dermaut [1] und Johnston [5], sodass hier auf weitere Auflistungen verzichtet werden kann.

Literatur

- ¹ Aelbers CMF, Dermaut LR. Orthopedics in orthodontics: Part I, Fiction or reality – a review of the literature. *Am J Orthod Dentofac Orthop* 1996; 110: 513 – 519
- ² Dermaut LR, Aelbers CMF. Orthopedics in Orthodontics: Part II, Fiction or reality – a review of the literature. *Am J Orthod Dentofac Orthop* 1996; 110: 667 – 671
- ³ Fränkel R, Fränkel C. Orofacial Orthopedics with the Function Regulator. Karger, Basel 1989; (Zitat S. 165)
- ⁴ Gianelly AA. One-phase versus two-phase treatment. *Am J Orthod Dentofac Orthop* 1995; 108: 556 – 559
- ⁵ Johnston LE. Growing jaws for fun and profit: a modest proposal. In: McNamara JA (Hrsg). Growth modification: what works, what doesn't, and why. Craniofacial Growth Series, Bd 35, Center for Human Growth and Development, The University of Michigan, Ann Arbor 1999; 63 – 86
- ⁶ Livieratos FA, Johnston LE. A comparison of one-stage and two-stage nonextraction alternatives in matched Class II samples. *Am J Orthod Dentofacial Orthop* 1995; 108: 118 – 131
- ⁷ Mills CM, McCulloch KJ. Posttreatment changes after successful correction of Class II malocclusions with the twin block appliance. *Am J Orthod Dentofacial Orthop* 2000; 118: 24 – 33
- ⁸ Proffit WR, Tulloch JF. Preadolescent Class II problems: treat now or wait? *Am J Orthod Dentofacial Orthop* 2002; 121: 560 – 562
- ⁹ Tulloch JF, Phillips C, Proffit WR. Benefit of early Class II treatment: progress report of a two-phase randomized clinical trial. *Am J Orthod Dentofacial Orthop* 1998; 113: 62 – 72